

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-514179

(P2004-514179A)

(43) 公表日 平成16年5月13日(2004.5.13)

(51) Int. Cl.<sup>7</sup>

F I

テーマコード (参考)

G10L 19/02

G10L 7/04

F

G10L 13/00

G10L 7/04

E

G10L 19/00

G10L 7/02

D

G10L 9/08

J

審査請求 有 予備審査請求 有 (全 43 頁)

(21) 出願番号 特願2002-543427(P2002-543427)  
 (86) (22) 出願日 平成13年11月13日(2001.11.13)  
 (85) 翻訳文提出日 平成15年5月12日(2003.5.12)  
 (86) 国際出願番号 PCT/SE2001/002510  
 (87) 国際公開番号 W02002/041301  
 (87) 国際公開日 平成14年5月23日(2002.5.23)  
 (31) 優先権主張番号 0004163-2  
 (32) 優先日 平成12年11月14日(2000.11.14)  
 (33) 優先権主張国 スウェーデン(SE)

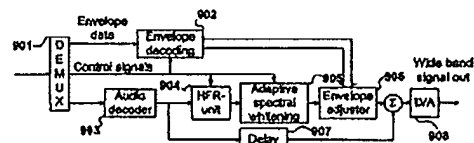
(71) 出願人 502112267  
 コーディング テクノロジーズ アクチボ  
 ラゲット  
 スウェーデン国 ストックホルム、ドベル  
 ンスガタン 64  
 (74) 代理人 100085464  
 弁理士 野口 繁雄  
 (72) 発明者 クリストファー・クジェルリング  
 スウェーデン国 S-17075 ゾルナ  
 ロスティゲン 10  
 (72) 発明者 ベル・エクストランド  
 スウェーデン国 S-11640 ストッ  
 クホルム ゼーデルマンナガタン 45

最終頁に続く

(54) 【発明の名称】 適応ろ波による高周波復元符号化方法の知覚性能の強化方法

(57) 【要約】

本発明は、高周波復元(HFR)を利用する音源符号化システムを強化する新しい方法及び装置を提案する。本発明は、適応ろ波を利用して、HFRが実行される音声信号の様々な周波数範囲の様々な音色特性によるアーチファクトを削減する。本発明は、発話符号化システム及び自然音符号化システムの両方に適用可能である。



**【特許請求の範囲】****【請求項 1】**

高周波復元を使用する音源符号化システムを強化する方法であって、前記音源符号化システムが記憶又は送信前に実行されるすべての演算を表すエンコーダと、記憶又は送信後に実行されるすべての演算を表すデコーダとを含んでいる方法において、前記エンコーダにおいて、任意の時間にオリジナル信号の音色特性を推定し、前記エンコーダにおいて、前記デコーダで H F R 方法が使用されるとき、前記デコーダでの H F R 後に類似する音色特性を得るために、任意の時間にスペクトル白色化の必要量を推定し、前記エンコーダから前記デコーダにスペクトル白色化の前記量に関する情報を送信し、  
前記デコーダにおいて、前記エンコーダから得られたスペクトル白色化情報に従って、高周波復元 (H F R) の前後に、適応可能に信号をスペクトル白色化することとを特徴とする方法。

10

**【請求項 2】**

オリジナル信号の音色特性の前記推定が、様々な周波数領域で行なわれることを特徴とする、請求項 1 記載の方法。

**【請求項 3】**

白色化の必要量の前記推定が、様々な周波数領域で行なわれることを特徴とする、請求項 1 記載の方法。

**【請求項 4】**

前記スペクトル白色化が、時間ドメインで実行されることを特徴とする、請求項 1 記載の方法。

20

**【請求項 5】**

前記スペクトル白色化が、サブバンドフィルタバンクで実行されることを特徴とする、請求項 1 記載の方法。

**【請求項 6】**

スペクトル白色化の必要量の前記推定が、前記オリジナル信号のサブバンドろ波から得られる様々なサブバンド信号の音色対雑音信号比  $q$  の比較により実行され、前記音色対雑音信号比が前記サブバンド信号の線形予測を使用して得られることを特徴とする、請求項 1 記載の方法。

30

**【請求項 7】**

スペクトル白色化の必要量の前記推定が、前記オリジナル信号及び H F R 信号のサブバンドろ波から得られる様々なサブバンド信号の音色対雑音信号比  $q$  の比較により実行され、前記音色対雑音信号比が前記サブバンド信号の線形予測を使用して得られ、前記 H F R 信号が前記デコーダにおける前記 H F R と同じ方法で生成されることを特徴とする、請求項 1 記載の方法。

**【請求項 8】**

スペクトル白色化の量が、L P C 予測子次数により制御されることを特徴とする、請求項 1 記載の方法。

**【請求項 9】**

スペクトル白色化の量が、L P C 多項式の帯域幅膨張係数により制御されることを特徴とする、請求項 1 記載の方法。

40

**【請求項 10】**

スペクトル白色化の量が、混合係数  $b$  により制御されることを特徴とする、請求項 1 記載の方法。

**【請求項 11】**

フィルタバンク解析フィルタの特性を補償するために、事前ろ波が L P C 推定に含まれることを特徴とする、請求項 5 記載の方法。

**【請求項 12】**

高周波復元を使用する音源符号化システムを強化する装置であって、前記音源符号化シス 50

テムが記憶又は送信前に実行されるすべての演算を表すエンコーダと、記憶又は送信後に実行されるすべての演算を表すデコーダとを含む装置において、前記エンコーダにおいて、任意の時間にオリジナル信号の音色特性を推定する手段と、前記エンコーダにおいて、前記デコーダでHFR方法が使用されるとしたとき、前記デコーダでのHFR後に類似する音色特性を得るために、任意の時間にスペクトル白色化の必要量を推定する手段と、前記デコーダにおいて、前記エンコーダから得られたスペクトル白色化情報に従って、高周波復元(HFR)の前後に、適応可能に信号をスペクトル白色化する手段とを備えることを特徴とする装置。

【発明の詳細な説明】

10

【0001】

(技術分野)

本発明は、スペクトル帯域複製、SBR [WO 98/57436号] 又は関連方法等の高周波復元(HFR)を利用する音源符号化システムに関する。本発明は、高質方法(SBR)並びに低質方法[米国特許第5,127,054号]の性能を改善する。本発明は、発話の符号化及び自然音の符号化共に適応可能である。

【0002】

(発明の背景)

高帯域が低帯域から外挿される音声信号の高周波復元において、HFRシステムで一般に使用されるような粗いエンベロープ調整で達成可能であるよりもさらに広い範囲まで復元された高帯域の音色成分を制御する手段を有することが重要である。これが必要な理由は、声等の多くの音声信号及び多くの音響機器の音色成分が、通常は、高周波領域と比較して低周波領域(すなわち4~5kHz以下)においてより強いためである。極端な例は、低帯域での非常に顕著な調和級数や、高帯域での多かれ少なかれ純粋な雑音である。これに取り組むための1つの方法は、再構築された高帯域に雑音を適応可能に追加することである(適応雑音追加[PCT/SE00/00159号])。しかしながら、これは低帯域の音色の特性を抑制するほど十分ではない場合があり、再構築された高帯域に反復性の「バズ音」特性を与える。さらに、雑音の正しい時間特性を達成することが困難である場合がある。一方が高い倍音密度(低ピッチ)をもち、他方が低い倍音密度(高ピッチ)をもつ2つの調和級数が混合されるときに別の問題が生じる。高ピッチの調和級数が低帯域では他方を支配するが、高帯域では支配しない場合、HFRは高ピッチ信号の高調波に高帯域を支配させ、復元された高帯域音をオリジナルに比べて「金属的」にする。前述のシナリオのいずれも、HFRシステムで一般的に使用されるエンベロープ調整を使用して制御することはできない。HFR信号のスペクトルエンベロープ調整の間に一定程度のスペクトル白色化が導入される実例もある。これは、その特定の程度のスペクトル白色化が所望の場合には満足の行く結果を与えるが、その特定の程度のスペクトル白色化から恩恵を受けない信号の抜粋には深刻なアーチファクトを生じさせる。

【0003】

(発明の概要)

本発明は、HFR方法に共通して生じる「バズ音」及び「金属音」の問題に関する。本発明は、エンコーダ側で高性能な検出アルゴリズムを使用して、デコーダで適用されるスペクトル白色化の好適な量を推定する。スペクトル白色化は、周波数でのみならず経時的にも変化し、複製される高帯域の高調波含有量を制御するための最善の手段を確実にする。本発明は、サブバンドフィルタバンク法においてだけでなく、時間ドメイン法においても実行可能である。

【0004】

本発明の特徴は以下の通りである。

エンコーダにおいて、任意の時間に様々な周波数領域のオリジナル信号の音色特性を推定する。

エンコーダにおいて、デコーダでHFR方法が使用されるとしたとき、デコーダでのHFR

50

R後に類似する音色特性を得るために、任意の時間に様々な周波数領域のスペクトル白色化の必要量を推定する。

エンコーダからデコーダへ、スペクトル白色化の好適な度合に関する情報を送信する。

デコーダにおいて、エンコーダから送信された情報に従って、時間ドメイン又はサブバンドフィルタバンクのいずれかでスペクトル白色化を実行する。

デコーダにおけるスペクトル白色化に使用される適応フィルタは、線形予測を使用して得られる。

必要とされるスペクトル白色化の度合は、予測によりエンコーダで査定される。

スペクトル白色化の度合は、予測子次数を変えることにより、LPC多項式の帯域幅膨張係数を変えることにより、あるいは任意の範囲までろ波された信号を未処理対照物と混合することにより制御される。 10

下位予測子を達成するサブバンドフィルタバンクを使用する能力は、特にフィルタバンクがすでにエンベロープ調整に使用されるシステムにおいて非常に効果的な実施を提供する。

スペクトル白色化の周波数選択度合は、本発明の新規フィルタバンクを実行すれば容易に得られる。

【0005】

ここで、添付図面を参照しながら本発明の範囲又は精神を制限することなく例示的な実施例によって本発明を説明する。

【0006】

20

(好適な実施形態の説明)

後述する実施形態は、高周波復元システムを改善するための本発明の原則について例示的であるにすぎない。ここに説明する装置及び詳細の変型並びに変化がこの技術に精通した他者に明らかになることが理解される。したがって、添付の特許クレームの範囲によってのみ制限され、実施形態の記述及び説明によって提示される特定の詳細によっては制限されないことを意図している。

【0007】

信号のスペクトルエンベロープを、任意のスペクトルエンベロープに調整すると、一定量のスペクトル白色化が常に適用される。これは、送信された粗いスペクトルエンベロープが $H_{envRef}(z)$ によって記述され、現在の信号セグメントのスペクトルエンベロープが $H_{envCur}(z)$ によって記述されるとすると、適用されるフィルタ関数は次式であるからである。 30

【数1】

$$W(z) = \frac{H_{envRef}(z)}{H_{envCur}(z)} \quad (1)$$

【0008】

本発明において、 $H_{envRef}(z)$ の周波数分解能は、 $H_{envCur}(z)$ の場合と必ずしも同じではない。本発明は、HFR信号のエンベロープ調整に $H_{envCur}(z)$ の適応周波数分解能を使用する。信号セグメントは、式1に従って信号をスペクトル白色化するために $H_{envCur}(z)$ の逆数でろ波される。 $H_{envCur}(z)$ は、線形予測を使用して得る場合、次式、 40

【数2】

$$H_{envCur}(z) = \frac{G}{A(z)} \quad (2)$$

によって記述される。式中、

【数3】

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k} \quad (3)$$

は、自己相関法又は共分散法〔「発話信号のデジタル処理」ラビナー&シェイファー、プレンティスホール社、ニュージャージー州、07632、イングルウッドクリフズ、ISBN 0-13-213603-1、第8章〕を使用して得られる多項式であり、Gは利得である。これを考慮すると、スペクトル白色化の度合は予測子次数を変える、すなわち多項式A(z)の次数を制限し、このようにしてH<sub>inv</sub>(Z)によって記述できる細かな構造の量を制限することにより、あるいは多項式A(z)に帯域幅膨張係数を適用することにより制御できる。帯域幅膨張は、以下に従って定義される。すなわち、帯域幅膨張係数をρとすれば、多項式A(z)は次式となる。

【数4】

$$A(\rho z) = a_0 z^0 \rho^0 + a_1 z^1 \rho^1 + a_2 z^2 \rho^2 + \dots + a_p z^p \rho^p. \quad (4)$$

【0009】

これは、図1に従ってH<sub>inv</sub>(Z)によって推定されるフォルマントの帯域幅を拡大する。指定時間の逆フィルタは、このようにして本発明に従って次式のように記述される。

【数5】

$$H_{inv}(z, \rho) = \frac{1 - \sum_{k=1}^p \alpha_k (z\rho)^{-k}}{G}, \quad (5)$$

式中、pは予測子次数であり、ρは帯域幅膨張係数である。

【0010】

係数α<sub>k</sub>は、前述のように、自己相関法又は共分散法等の様々な方法で得られる。利得係数Gは、定期的なエンベロープ調整の前にH<sub>inv</sub>が使用されると1に設定できる。システムの安定性を保証するために、なんらかの種類の緩和を推定値に追加することは一般的な慣習である。自己相関法を使用すると、これは相関ベクタのゼロ遅延値を相殺することによって容易に達成される。これは、A(z)を推定するために使用した信号に対する、一定レベルの白雑音の追加と同等である。パラメータpとρは、エンコーダから送信された情報に基づいて計算される。

【0011】

帯域幅膨張の代替策は、

【数6】

$$A_b(z) = 1 - b + b \cdot A(z), \quad (6)$$

40

によって記述される。式中、bは混合係数である。これは、次式に従って適応フィルタを生成する。

【数7】

$$H_{inv}(z, p, b) = \frac{1 - b + b \cdot \left( 1 - \sum_{k=1}^p \alpha_k (z)^{-k} \right)}{G}. \quad (7)$$

【0012】

50

式中、 $b = 1$ の場合、 $\rho = 1$ で数式7が数式5となり、 $b = 0$ の場合数式7が一定の非周波数選択利得係数となることが明らかである。

#### 【0013】

デコーダで使用される白色化の程度に関する情報は非常に効率的に送信できるため、本発明は、非常に低い追加ビットレートコストでHFRシステムの性能を劇的に高める。図2～図4は、例示的な絶対スペクトルによって、本発明を使用するシステムの性能を、本発明を使用しないシステムと比較して表示するものである。図2において、時間 $t_0$ 及び時間 $t_1$ でのオリジナル信号の絶対スペクトルを表示す。信号の低帯域及び高帯域の音色特性は時間 $t_0$ では類似しているが、時間 $t_1$ でそれらが大幅に異なることは明らかである。図3において、本発明を使用しないコピーアップベースのHFRを使用するシステムの時間 $t_0$ 及び時間 $t_1$ での出力を示す。ここでは、スペクトル白色化は適用されず、時間 $t_0$ で正しい音色特性、時間 $t_1$ では完全に誤った音色特性を与える。これにより非常にうるさいアーチファクトが生じる。アーチファクトは異なる特性を有し、様々な段階で発生するが、任意の一定した度合のスペクトル白色化について類似した結果が得られる。図4において、本発明を使用するシステムの時間 $t_0$ 及び時間 $t_1$ での出力を示す。ここでは、スペクトル白色化の量が経時的に変化し、その結果本発明を使用しないシステムの音質よりはるかに優れた音質となることが明らかである。

#### 【0014】

(エンコーダ側の検出器)

本発明では、現在使用されているHFR方法を考慮して、可能な限りオリジナルに類似した高帯域を得るために、デコーダで使用されるスペクトル白色化(LPC次数、帯域幅膨張係数及び/又は混合係数)の最良の度合を査定するためにエンコーダ側の検出器を使用する。デコーダで使用されるスペクトル白色化の度合の適切な推定値を得るにはいくつかの方法が使用可能である。後述する説明では、HFRアルゴリズムは、高周波の生成の間に低帯域スペクトルの音色構造を実質的には改變しない、つまり生成された高帯域は低帯域と同じ音色の特性を備えると仮定する。このような仮定を立てることができない場合、以下の検出は合成による解析を使用して、つまりエンコーダでオリジナル信号にHFRを実行することによって実行でき、オリジナル信号の低帯域及び高帯域に関する比較研究を実行するより、2つの信号の高帯域に関する比較研究を実行することができる。

#### 【0015】

1つの方法は、スペクトル白色化の適切な量を推定するために自己相関を使用する。検出器は、ソース範囲(すなわち、HFRがデコーダ内で基にする周波数範囲)及びターゲット範囲(すなわち、デコーダ内で復元される周波数範囲)の自己相関関数を推定する。図5aに、低帯域の調和級数及び高帯域の白色雑音をもつ、最悪の場合の信号を記述する。図5bに、様々な自己補正関数を示す。ここで、低帯域はきわめて相互に関連付けられるが、高帯域は相互に関連付けられていないことが明らかである。最小の遅延より大きな遅延について最大相関が高帯域及び低帯域に得られる。この2つの比率が、デコーダで適用されるスペクトル白色化の最適な度合を算出するために使用される。本発明を前述において概略したように実現すると、相関の算出にFFTを使用することが好適である場合がある。数列 $x(n)$ の自己相関は次式により定義される。

#### 【数8】

$$r_{xx}(m) = FFT^{-1}(|X(k)|^2), \quad (8)$$

式中、

#### 【数9】

$$X(k) = FFT(x(n)). \quad (9)$$

である。

【0016】

目的は高帯域及び低帯域での自己相関の差異を比較することであるため、ろ波は周波数ドメインで実行できる。これにより、

【数10】

$$\begin{cases} X_{Lp}(k) = X(k) \cdot H_{Lp}(k) \\ X_{Hp}(k) = X(k) \cdot H_{Hp}(k) \end{cases} \quad (10)$$

10

が得られる。式中、 $H_{Lp}(k)$  及び  $H_{Hp}(k)$  はLP及びHPフィルタインパルス応答のフーリエ変換である。

【0017】

前述から、低帯域及び高帯域の自己相関関数を、次式に従って算出することができる。

【数11】

$$\begin{cases} r_{xxLp}(m) = FFT^{-1}(|X_{Lp}(k)|^2) \\ r_{xxHp}(m) = FFT^{-1}(|X_{Hp}(k)|^2) \end{cases} \quad (11)$$

20

【0018】

自己相関ベクタごとに最小遅延より大きな遅延について最大値が算出される。

【数12】

$$\begin{cases} r_{MaxLp} = \max(r_{xxLp}) \quad \forall m > \minLag \\ r_{MaxHp} = \max(r_{xxHp}) \quad \forall m > \minLag \end{cases} \quad (12)$$

30

【0019】

この2つの定数は、適切な帯域幅膨張係数に例えばマッピングするために使用できる。

【0020】

前述の説明は、任意の時間での任意の周波数帯域の正しい逆ろ波レベルを取得するために、予測性、すなわち任意の時間における任意の周波数帯域での信号の音色対雑音比、の一般的な測定値を査定することが有益であることを暗示している。これは、以下のさらに精緻化された方法を使用して達成できる。ここでは、サブバンドフィルタバンクが仮定されるが、本発明がこのようなものに制限されないことが十分に理解される。

40

【0021】

フィルタバンクの各サブバンドの音色対雑音比  $q$  は、サブバンドサンプルのブロックで線形予測を使用することによって定義できる。 $q$  の大きな値は大量の調性を示すが、 $q$  の小さな値はその信号が時間及び周波数に関し、対応場所での雑音状であることを示す。 $q$  値は共分散法と自己相関法の両方を使用して得られる。

【0022】

共分散法について、サブバンド信号ブロック  $[x(0), x(1), \dots, x(N-1)]$  の線形予測係数及び予測誤差を、Cholesky分解[「発話信号のデジタル処理」ラビナー&シェイファー、プレントイスホール社、ニュージャージー州、07632、イングルウッドクリフズ、ISBN 0-13-213603-1、第8章]を使用して

50

効率的に算出できる。次に、音色対雑音比  $q$  が次式によって定義される。

【数 1 3】

$$q = \frac{\Psi - E}{E}, \quad (13)$$

式中、 $\Psi = |x(0)|^2 + |x(1)|^2 + \dots + |x(N-1)|^2$  は信号ブロックのエネルギーであり、 $E$  は予測誤差ブロックのエネルギーである。

【0023】

自己相関法の場合、さらに自然な方法は、Levinson-Durbin アルゴリズム 10  
 [「デジタル信号処理、原則、アルゴリズム及び応用例」、第3版、John G. Proakis, Dimitris G. Manolakis、プレントイスホール、国際版、ISBN-0-13-394338-9、第11章] を使用することであり、その場合、 $q$  は次式に従って定義される。

【数 1 4】

$$q = \left( \prod_{i=1}^p (1 - |K_i|^2) \right)^{-1} - 1, \quad (14)$$

式中、 $K_i$  は予測多項式から得られる対応する格子フィルタ構造の反射係数であり、 $p$  は 20  
 予測子次数である。

【0024】

次に、復元された高帯域の音色対雑音比がオリジナル高帯域の音色対雑音比に近づくようにスペクトル白色化の度合を調整するために、 $q$  の高帯域値と低帯域値の比率が使用される。ここでは、混合係数  $b$  (数式 6) を利用して白色化の程度を制御するのが有利である。

【0025】

音色対雑音比  $q = q_H$  が高帯域で測定され、 $q = q_L \geq q_H$  が低帯域で測定されると仮定すると、白色化係数  $b$  の適切な選択肢が以下の式で与えられる。

【数 1 5】

$$b = 1 - \sqrt{\frac{q_H}{q_L}}. \quad (15)$$

【0026】

これを確かめるために、第1ステップは、数式 6 を以下の形式に書き直すことである。

【数 1 6】

$$A_b(z) = A(z) + (1-b)(1-A(z)). \quad (16)$$

【0027】

これは、 $A(z)$  を推定するために使用された信号が、フィルタ  $A_b(z)$  によりろ波される場合、予測された信号が利得係数  $1-b$  分抑制され、予測誤差が改変されないことを示している。音色対雑音比は、平均二乗予測信号対平均二乗予測誤差の比率であるため、ろ波前の  $q$  の値は、ろ波動作により  $(1-b)^2 q$  に変更される。これを低帯域信号に適用すると、音色対雑音比  $(1-b)^2 q$  の信号が生成され、適用された HFR 法が調性を改変しないと仮定して、 $b$  が正確に数式 15 に従って選択されるのであれば、高帯域のターゲット値  $q_H$  に達する。

30

40

50



## 【0028】

64チャネルフィルタバンクの各サブバンドでの予測次数  $p = 2$  に基づく  $q$  の値は、図5aの信号について図5cに示す。騒々しい部分より信号の高調波部分について大幅に高い値に達する。高調波部分での推定値の可変性は、選ばれる周波数分解能及び予測次数による。

## 【0029】

(時間ドメインにおける適応LPCベースの白色化)

デコーダでの適応ろ波は、高周波復元の前後に実行できる。ろ波がHFRの前に実行される場合、使用されるHFR法の特性を考慮する必要がある。周波数選択適応ろ波が実行される場合は、低帯域領域に正しい量のスペクトル白色化を適用するために、システムは、HFR部の前で、特定の高帯域領域がその低帯域領域を生じるものから控除しなければならない。本発明の時間ドメイン法の以下の例では、非周波数選択適応スペクトル白色化が概略的に説明される。本発明の時間ドメイン法が後述の実施例に制限されないことはこの技術に精通した者にとって明らかである。

## 【0030】

時間ドメインで適応ろ波を実行する際は、自己相関法を使用する線形予測が好ましい。自己相関法は、係数  $\alpha_k$  を推定するために使用される入力セグメントのウィンドウ処理を必要とするが、このことは共分散法には当てはまらない。本発明によるスペクトル白色化に使用されるフィルタは、

## 【数17】

$$H_{inv}(z, p, \rho) = 1 - \sum_{k=1}^p \alpha_k (z\rho)^{-k}, \quad (19)$$

であり、式中、利得係数  $G$  (数式5) が1に設定される。適応スペクトル白色化がHFR部の前に実行される場合、適応フィルタはさらに低いサンプリング速度で動作できるため、効果的な実行が達成される。低帯域信号は、図6に従って、エンコーダにより与えられた予測子次数及び帯域幅膨張係数をもってウィンドウ処理され、適切な時間ベースでろ波される。本発明のこの実施例では、信号は低域ろ波され(601)、大量に除去される(602)。603は適応フィルタを示す。ウィンドウ606は  $A(z)$  多項式の推定のための適切な時間セグメントを選択するために使用され、50%の重複が使用される。LPCルーチン607は、一般的に選ばれたLPC次数及び帯域幅膨張係数が与えられた  $A(z)$  を、適切な緩和をもって抽出する。FIRフィルタ608は、信号セグメントを適応可能にろ波するために使用される。スペクトル白色化された信号セグメントはアップサンプルされ(604、605)、共にウィンドウ処理され、HFR部に対する入力信号を形成する。

## 【0031】

(サブバンドフィルタバンクでの適応LPCベースの白色化)

適応ろ波は、フィルタバンクを使用して効果的かつ堅牢に実行できる。線形予測及びろ波は、フィルタバンクにより発生するサブバンド信号のそれぞれに、独立して実行される。サブバンド信号のエイリアス成分が抑制されるフィルタバンクを使用することが有利である。これは、例えばフィルタバンクをオーバーサンプルすることによって達成できる。例えば適応ろ波の結果生じる、サブバンド信号の個々の変型から出現するエイリアスのためのアーチファクトは、大きく削減できる。サブバンド信号のスペクトル白色化は、前述された時間ドメイン法に類似する線形予測を通して取得される。サブバンド信号が複素値である場合、ろ波だけではなく線形予測にも複雑なフィルタ係数が使用される。各周波数バンドの音色成分の予想数は妥当な量のフィルタバンクチャネルのあるシステムに対しては非常に小さいため、線形予測の次数を非常に低く保つことができる。時間ドメインLPCと同じ時間ベースに対応するために、各ブロックのサブバンドサンプル数はフィルタバンクのダウンサンプルに等しい係数分小さくなる。低フィルタ次数及び小さなブロックサイズ

を考慮すると、予測フィルタ係数は、好ましくは共分散法を使用して得られる。フィルタ係数の算出及びスペクトル白色化は、ブロック長 $N$ より小さいサブバンドサンプルタイムステップ $L$ を使用してブロック単位で実行できる。スペクトル白色化されたブロックは、適切な合成ウィンドウ処理を使用してともに追加されなければならない。

#### 【0032】

最大限に大量に削減されたフィルタバンクに、白ガウス雑音からなる入力信号を与えると、白色スペクトル密度のサブバンド信号が生成される。オーバーサンプルフィルタバンクに白色雑音を与えると、サブバンド信号に着色されたスペクトル密度が与えられる。これは、解析フィルタの周波数応答の影響によるものである。雑音状の入力信号の場合、フィルタバンクチャネルの $LPC$ 予測子はフィルタ特性を追跡する。これは求められていない特徴であり、補償の恩恵を受ける。考えられる解決策は、線形予測子に対する入力信号の事前ろ波である。事前ろ波は、解析フィルタの周波数応答を補償するために、解析フィルタの逆数又は逆数の近似でなければならない。白色化フィルタには、前述のようにオリジナルサブバンド信号が与えられる。図7は、サブバンド信号の白色化工程を示す。チャネル1に対応するサブバンド信号は事前ろ波ブロック701に送られ、その後、深度がフィルタ次数702に依存する遅延チェーンに送られる。遅延した信号及びそれらの共役703が線形予測ブロック704に送られ、係数が算出される。 $L$ 番目の各算出の係数は、デシメータ705によって維持される。サブバンド信号は最終的に、予測される係数が $L$ 番目のサンプルごとに使用、更新されるフィルタブロック706を通ってろ波される。

#### 【0033】

##### (実践的な実施)

本発明は、任意のコーデックを使用して、アナログもしくはデジタル信号の記憶又は送信のために多様な種類のシステムについてハードウェアチップとDSPの両方で実現できる。図8及び図9は、本発明の考えられる実施例を示す。図8はエンコーダ側を示す。アナログ入力信号は、逆ろ波レベル推定部803及びエンベロープ抽出部804だけではなく、 $A/D$ 変換器801及び任意の音声コーデ802にも送られる。コード化された情報は連続的なビットストリーム805に多重化され、送信又は記憶される。図9は、典型的なデコーダ実施例を示す。連続的なビットストリームは非多重化(901)され、エンベロープデータ、すなわち高帯域のスペクトルエンベロープは復号(902)される。非多重化された符号化信号は任意の音声デコーダ903を使用して復号される。復号された信号は任意のHFR部904に送られ、そこで高帯域が再生される。高帯域信号は、適応スペクトル白色化を実行するスペクトル白色化装置905に送られる。その後、信号はエンベロープ調整器906に送られる。エンベロープ調整器からの出力は、遅延907を通して送られた復号された信号と結合される。最後に、デジタル出力がアナログ波形908に変換し直される。

##### 【図面の簡単な説明】

【図1】  $LPC$ スペクトルの帯域幅膨張を示す。

【図2】 時間 $t_0$ 及び時間 $t_1$ でのオリジナル信号の絶対スペクトルを示す。

【図3】 適応ろ波を使用しない、従来技術によるコピーアップHFRシステムの時間 $t_0$ 及び時間 $t_1$ での出力の絶対スペクトルを示す。

【図4】 本発明に係る、適応ろ波を使用するコピーアップHFRシステムの時間 $t_0$ 及び時間 $t_1$ での出力の絶対スペクトルを示す。

【図5a】 本発明に係る最悪の場合の信号を示す。

【図5b】 最悪の場合における信号の高帯域と低帯域の自己相関を示す。

【図5c】 本発明に係る様々な周波数の音色対雑音比 $q$ を示す。

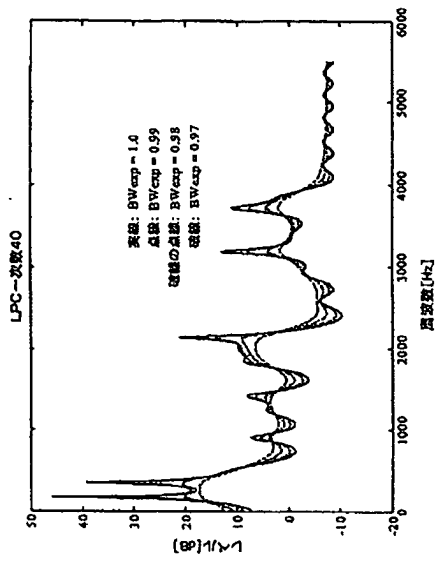
【図6】 本発明に係るデコーダでの適応ろ波の時間ドメイン法を示す。

【図7】 本発明に係るデコーダでの適応ろ波のサブバンドフィルバンク法を示す。

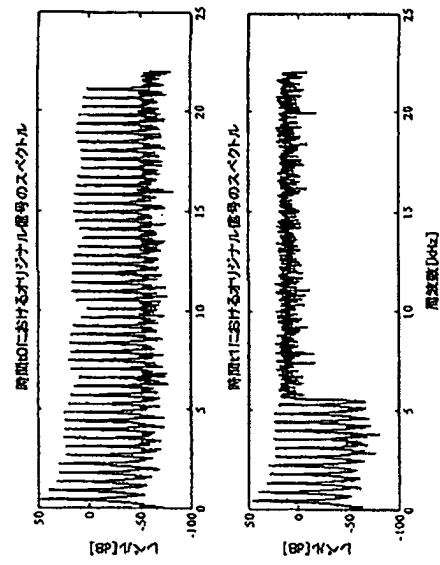
【図8】 本発明のエンコーダの実施例を示す。

【図9】 本発明のデコーダの実施例を示す。

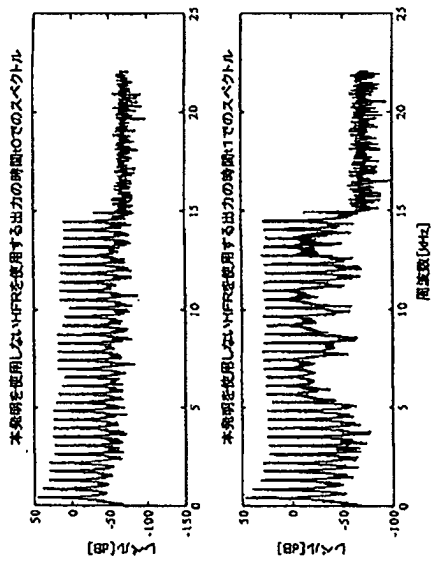
【図 1】



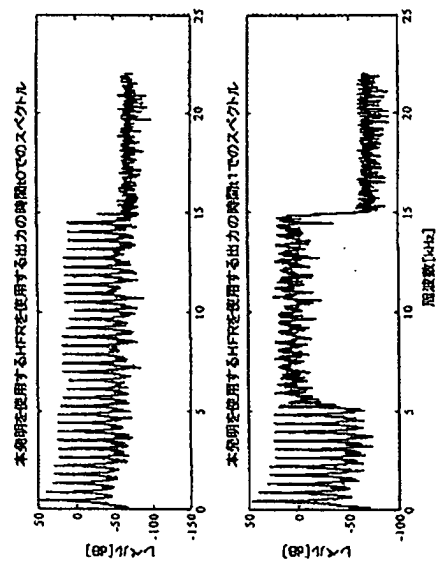
【図 2】



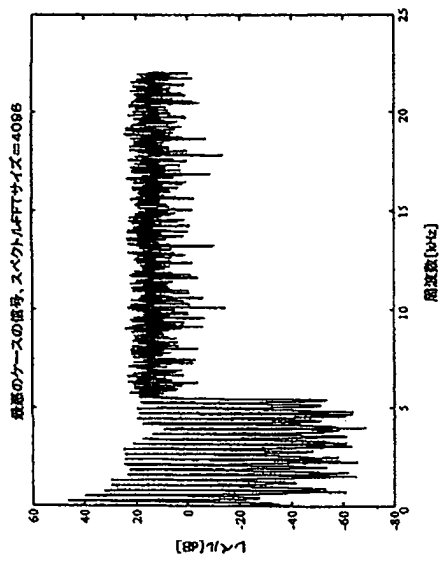
【図 3】



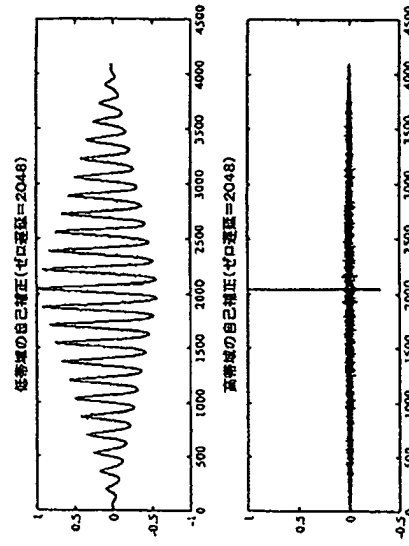
【図 4】



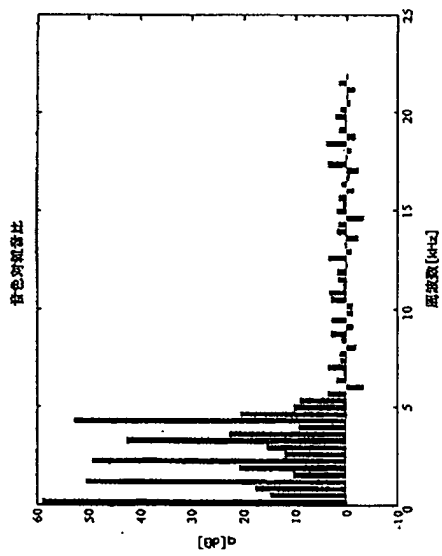
【図 5 a】



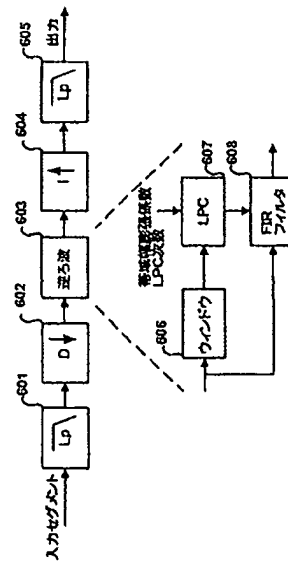
【図 5 b】



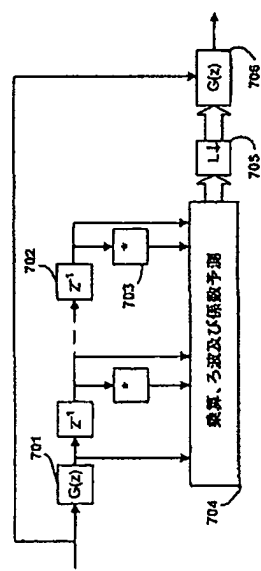
【図 5 c】



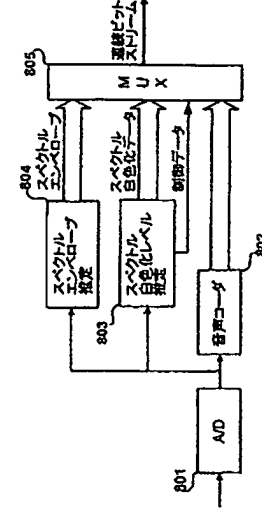
【図 6】



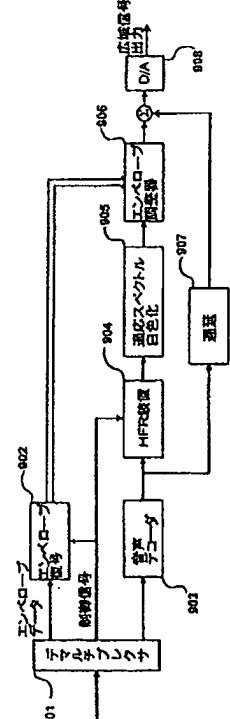
【図 7】



【図 8】



【図 9】



## 【国際公開パンフレット】

(1) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
23 May 2002 (23.05.2002)

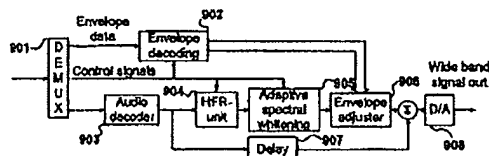
PCT

(10) International Publication Number  
WO 02/41301 A1

- (51) International Patent Classification: G10L 19/00 (74) Agents: ÖRTENGREN AD, Berzelij väg, Kungälv, 415 01 Göteborg (SE).
- (11) International Application Number: PCT/SE01/00210 (81) Designated States (national): AR, AO, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, ES, EA, EC, EG, FR, GB, GR, GT, HK, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MY, NZ, PA, PE, PG, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (12) International Filing Date: 13 November 2001 (13.11.2001) (15) Piling Language: English (16) Publication Language: English
- (13) Priority Date: 14 November 2000 (14.11.2000) SE (84) Designated States (epitaph): AR, AO, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, ES, EA, EC, EG, FR, GB, GR, GT, HK, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MY, NZ, PA, PE, PG, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (17) Applicant (for all designated States except US): CODING TECHNOLOGIES SWEDEN AB (SESE), Doktorsgatan 44, S-113 32 Stockholm (SE).
- (72) Inventors and (73) International Applicants (for US only): EKHOLM, Mats (SESE), Långgatan 10, S-170 75 Solna (SE); EKSTRAND, Per (SESE), Sköndalsvägen 45, S-141 40 Stockholm (SE); HEIN, Fredrik (SESE), Rindögatan 14, S-168 31 Bromma (SE); VILHELMSSON, Lars (SESE), Mandostråden 22, S-171 56 Rindögatan (SE).
- Published: with international search report
- For two-letter codes and other abbreviations, refer to the "Guide to Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ENHANCED PERCEPTUAL PERFORMANCE OF HIGH FREQUENCY RECONSTRUCTION CODING METHODS BY ADAPTIVE FILTERING

WO 02/41301 A1



(57) Abstract: The present invention proposes a new method and a new apparatus for enhancement of audio source coding systems using high frequency reconstruction (HFR). It utilizes adaptive filtering to reduce artifacts due to different input characteristics in different frequency ranges of an audio signal upon which HFR is performed. The present invention is applicable in both speech coding and musical audio coding systems.

# ENHANCING PERCEPTUAL PERFORMANCE OF HIGH FREQUENCY RECONSTRUCTION CODING METHODS BY ADAPTIVE FILTERING

## TECHNICAL FIELD

5 The present invention relates to audio source coding systems utilizing high frequency reconstruction (HFR) such as Spectral Band Replication, SBR [WO 98/57436] or related methods. It improves performance of high quality methods (SBR), as well as low quality methods (U.S. Pat. 5,127,054). It is applicable to both speech coding and natural audio coding systems.

## 10 BACKGROUND OF THE INVENTION

In high frequency reconstruction of audio signals, where a highband is extrapolated from a lowband, it is important to have means to control the tonal components of the reconstructed highband to a greater extent than what can be achieved with a coarse envelope adjustment, as commonly used in HFR systems. This is necessary since the tonal components for most audio signals such as voices and most acoustic instruments, usually are stronger in the low frequency regions (i.e. below 4-kHz) compared to the high frequency regions. An extreme example is a very pronounced harmonic series in the lowband and more or less pure noise in the high band. One way to approach this is by adding noise adaptively to the reconstructed highband (Adaptive Noise Addition [PCT/SE00/00159]). However, this is sometimes not enough to suppress the tonal character of the lowband, giving the reconstructed highband a repetitive "buzzy" sound character. Furthermore, it can be difficult to achieve the correct temporal characteristic of the noise. Another problem occurs when two harmonic series are mixed, one with high harmonic density (low pitch) and the other with low harmonic density (high pitch). If the high-pitched harmonic series dominates over the other in the lowband but not in the highband, the HFR causes the harmonics of the high-pitched signal to dominate the highband, making the reconstructed highband sound "metallic" compared to the original. None of the above-described scenarios can be controlled using the envelope adjustment commonly used in HFR systems. In some implementations a constant degree of spectral whitening is introduced during the spectral envelope adjustment of the HFR signal. This gives satisfactory results when that particular degree of spectral whitening is desired, but introduces severe artifacts for signal excerpts that do not benefit from that particular degree of spectral whitening.

## SUMMARY OF THE INVENTION

The present invention relates to the problem of "buziness" and "metallic"-sound that is commonly introduced in HFR-methods. It uses a sophisticated detection algorithm on the encoder side to estimate the preferable amount of spectral whitening to be applied in the decoder. The spectral whitening varies over time as well as over frequency, ensuring the best means to control the harmonic

contents of the replicated highband. The present invention can be carried out in a time-domain implementation as well as in a subband filterbank implementation.

The present invention comprises the following features:

- 5 - In the encoder, estimating the tonal character of an original signal for different frequency regions at a given time.
- In the encoder, estimating the required amount of spectral whitening, for different frequency regions at a given time, in order to obtain a similar tonal character after HFR in the decoder, given the HFR-method used in the decoder.
- 10 - Transmitting the information on preferred degree of spectral whitening from the encoder to the decoder.
- In the decoder, perform spectral whitening in either the time domain or in a subband filterbank, in accordance with the information transmitted from the encoder.
- The adaptive filter used for spectral whitening in the decoder is obtained using linear prediction.
- 15 - The degree of spectral whitening required is assessed in the encoder by means of prediction.
- The degree of spectral whitening is controlled by varying the predictor order, or by varying the bandwidth expansion factor of the LPC polynomial, or by mixing the filtered signal, to a given extent, with the unprocessed counterpart.
- The ability to use a subband filterbank achieving low-order predictors, offers very effective implementation, especially in a system where a filterbank already is used for envelope adjustment.
- 20 - Frequency selective degree of spectral whitening is easily obtained given the novel filterbank implementation of the present invention.

## 25 BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described by way of illustrative examples, not limiting the scope or spirit of the invention, with reference to the accompanying drawings, in which:

Fig. 1 illustrates bandwidth expansion of an LPC spectrum;

Fig. 2 illustrates the absolute spectrum of an original signal at time  $t_0$ , and time  $t_1$ ;

- 30 Fig. 3 illustrates the absolute spectrum of the output, at time  $t_0$  and time  $t_1$ , of a prior art copy up HFR system without adaptive filtering;

Fig. 4 illustrates the absolute spectrum of the output, at time  $t_0$  and time  $t_1$ , of a copy up HFR system with adaptive filtering, according to the present invention;

Fig. 5a illustrates a worst case signal according to the present invention;

- 35 Fig. 5b illustrates the autocorrelation for the highband and lowband of the worst case signal;

Fig. 5c illustrates the tonal to noise ratio  $q$  for different frequencies, according to the present invention;

Fig. 6 illustrates a time domain implementation of the adaptive filtering in the decoder, according to the present invention;

- 40 Fig. 7 illustrates a subband filterbank implementation of the adaptive filtering in the decoder, according to the present invention;



Fig. 8 illustrates an encoder implementation of the present invention;  
 Fig. 9 illustrates a decoder implementation of the present invention.

#### DESCRIPTION OF PREFERRED EMBODIMENTS

- 5 The below-described embodiments are merely illustrative for the principles of the present invention for improvement of high frequency reconstruction systems. It is understood that modifications and variations of the arrangements and the details described herein will be apparent to others skilled in the art. It is the intent, therefore, to be limited only by the scope of the impending patent claims and not by the specific details presented by way of description and explanation of the embodiments herein.
- 10 When adjusting a spectral envelope of a signal to a given spectral envelope a certain amount of spectral whitening is always applied. This, since if the transmitted coarse spectral envelope is described by  $H_{envRef}(z)$  and the spectral envelope of the current signal segment is described by  $H_{envCur}(z)$ , the filter function applied is

$$15. \quad W(z) = \frac{H_{envRef}(z)}{H_{envCur}(z)}. \quad (1)$$

- In the present invention the frequency resolution for  $H_{envRef}(z)$  is not necessarily the same as for  $H_{envCur}(z)$ . The invention uses adaptive frequency resolution of  $H_{envCur}(z)$  for envelope adjustment of BPF signals. The signal segment is filtered with the inverse of  $H_{envCur}(z)$ , in order to
- 20 spectrally whiten the signal according to Eq. 1. If  $H_{envCur}(z)$  is obtained using linear prediction, it can be described according to

$$H_{envCur}(z) = \frac{Q}{A(z)}, \quad (2)$$

where

$$A(z) = 1 - \sum_{i=1}^P \alpha_i z^{-i} \quad (3)$$

- 25 is the polynomial obtained using the autocorrelation method or the covariance method [Digital Processing of Speech Signals, Rabiner & Schafer, Prentice Hall, Inc., Englewood Cliffs, New Jersey 07632, ISBN 0-13-213603-1, Chapter 8], and Q is the gain. Given this, the degree of spectral whitening can be controlled by varying the predictor order, i.e. limiting the order of the polynomial  $A(z)$ , and thus limiting the amount of fine structure that can be described by  $H_{envCur}(z)$ , or by
- 30 applying a bandwidth expansion factor to the polynomial  $A(z)$ . The bandwidth expansion is defined according to the following; if the bandwidth expansion factor is  $\rho$ , the polynomial  $A(z)$  evaluates to

$$A(\rho z) = a_0 z^P \rho^P + a_1 z^{P-1} \rho^{P-1} + a_2 z^{P-2} \rho^{P-2} + \dots + a_P z^0 \rho^0. \quad (4)$$

This expands the bandwidth of the formants estimated by  $H_{\text{form}}(x)$  according to Fig. 1. The inverse filter at a given time is Qmf, according to the present invention, described as

$$H_{\text{inv}}(z, p, \rho) = \frac{1 - \sum_{k=1}^p \alpha_k (z\rho)^{-k}}{G}, \quad (5)$$

where  $p$  is the predictor order and  $\rho$  is the bandwidth expansion factor.

5 The coefficients  $\alpha_k$  can, as mentioned above, be obtained in different manners, e.g. the autocorrelation method or the covariance method. The gain factor  $G$  can be set to one if  $H_{\text{form}}$  is used prior to a regular envelope adjustment. It is common practice to add some sort of relaxation to the estimate in order to ensure stability of the system. When using the autocorrelation method this is easily accomplished by offsetting the zero-lag value of the correlation vector. This is equivalent to  
10 addition of white noise at a constant level to the signal used to estimate  $A(x)$ . The parameters  $p$  and  $\rho$  are calculated based on information transmitted from the encoder.

An alternative to bandwidth expansion is described by:

$$A_b(x) = 1 - b + b \cdot A(x), \quad (6)$$

where  $b$  is the blending factor. This yields the adaptive filter according to:

$$H_{\text{inv}}(z, p, b) = \frac{1 - b + b \cdot \left(1 - \sum_{k=1}^p \alpha_k (z\rho)^{-k}\right)}{G}. \quad (7)$$

Here it is evident that for  $b = 1$  Eq. 7 evaluates to Eq. 5 with  $\rho = 1$ , and for  $b = 0$  Eq. 7 evaluates to a constant non-frequency selective gain factor.

20 The present invention drastically increases the performance of HFR systems, at a very low additional bitrate cost, since the information on the degree of whitening to be used in the decoder can be transmitted very efficiently. Fig. 2 - 4 displays the performance of a system with the present invention compared to a system without, by means of illustrative absolute spectra. In Fig. 2 absolute spectra of the original signal at time  $t_0$  and time  $t_1$  are displayed. It is evident that the tonal character for the  
25 lowband and the highband of the signal is similar at time  $t_0$ , while they differ significantly at time  $t_1$ . In Fig. 3 the output at time  $t_0$  and time  $t_1$  of a system using a copy-op based HFR without the present invention are displayed. Here, no spectral whitening is applied giving the correct tonal character at time  $t_0$ , but entirely wrong at time  $t_1$ . This causes very annoying artifacts. Similar results would be  
30 obtained for any constant degree of spectral whitening, albeit the artifacts would have different characters and occur at different instances. In Fig. 4 the output at time  $t_0$  and time  $t_1$  of a system using

the present invention are displayed. Here it is evident that the amount of spectral whitening varies over time, which results in a sound quality far superior to that of a system without the present invention.

#### The detector on the encoder side

- 5 In the present invention, a detector on the encoder-side is used to assess the best degree of spectral whitening (LPC order, bandwidth expansion factor and/or bleading factor) to be used in the decoder, in order to obtain a highband as similar to the original as possible, given the currently used HFR method. Several approaches can be used in order to obtain a proper estimate of the degree of spectral whitening to be used in the decoder. In the following description below, it is assumed that the HFR  
 10 algorithm does not substantially alter the total structure of the lowband spectrum during the generation of high frequencies, i.e. the generated highband has the same tonal character as the lowband. If such assumptions cannot be made the below detection can be performed using an analysis by synthesis, i.e. performing HFR on the original signal in the encoder and do the comparative study on the highbands of the two signals, rather than doing a comparative study on the lowband and highband of the original signal.

- One approach uses autocorrelation to estimate the appropriate amount of spectral whitening. The detector estimates the autocorrelation functions for the source range (i.e. the frequency range upon which the HFR will be based in the decoder) and the target range (i.e. the frequency range to be  
 20 reconstructed in the decoder). In Fig 5a, a worst case signal is described, with a harmonic series in the lowband and white noise in the highband. The different autocorrelation functions are displayed in Fig 5b. Here it is evident that the lowband is highly correlated whilst the highband is not. The maximum correlation, for any lag larger than a minimum lag, is obtained for both the highband and the lowband. The quotient of the two is used to calculate the optimal degree of spectral whitening to be applied in the decoder. When implementing the present invention as outlined above, it may be preferable to use FFTs for the computation of the correlation. The autocorrelation of a sequence  $x(n)$  is defined by:

$$r_{xx}(m) = FFT^{-1}(|X(k)|^2), \quad (8)$$

where

$$X(k) = FFT(x(n)). \quad (9)$$

Since the objective is to compare the difference of the autocorrelation in the highband and the lowband the filtering can be done in the frequency domain. This yields:

$$\begin{cases} X_{LP}(k) = X(k) \cdot H_{LP}(k) \\ X_{HP}(k) = X(k) \cdot H_{HP}(k) \end{cases}, \quad (10)$$

where  $H_{LP}(k)$  and  $H_{HP}(k)$  are the Fourier transforms of the LP and HP filters impulse responses.

From the above the autocorrelation functions for the lowband and highband can be calculated according to:

$$\begin{cases} r_{\text{mlp}}(m) = FFT^{-1} \left( \left| X_{\text{lp}}(k) \right|^2 \right) \\ r_{\text{hlp}}(m) = FFT^{-1} \left( \left| X_{\text{hp}}(k) \right|^2 \right) \end{cases} \quad (11)$$

The maximum value, for a lag larger than a minimum lag, for each autocorrelation vector is calculated:

$$\begin{cases} r_{\text{mlp}} = \max(r_{\text{mlp}}) \quad \forall m > \text{minlag} \\ r_{\text{hlp}} = \max(r_{\text{hlp}}) \quad \forall m > \text{minlag} \end{cases} \quad (12)$$

The quota of the two can be used to for instance map to a suitable bandwidth expansion factor.

The above implies that it would be beneficial to assess a general measurement of the predictability, i.e. the tonal to noise ratio of a signal in a given frequency band at a given time, in order to obtain a correct inverse filtering level for a given frequency band at a given time. This can be accomplished using the more refined approach below. Here a subband filterbank is assumed, it is well understood however that the invention is not limited to such.

A tonal to noise ratio  $q$  for each subband of a filter bank can be defined by using linear prediction on blocks of subband samples. A large value of  $q$  indicates a large amount of tonality, whereas a small value of  $q$  indicates that the signal is noise-like at the corresponding location in time and frequency. The  $q$ -value can be obtained using both the covariance method and the autocorrelation method.

For the covariance method, the linear prediction coefficients and the prediction error for the subband signal block  $[x(0), x(1), \dots, x(N-1)]$  can be computed efficiently by using the Cholesky decomposition, [Digital Processing of Speech Signals, Rabiner & Scherer, Prentice Hall, Inc., Englewood Cliffs, New Jersey 07632, ISBN 0-13-213603-1, Chapter 8]. The tonal to noise ratio  $q$  is then defined by

$$q = \frac{\Psi - E}{E}, \quad (13)$$

where  $\Psi = |x(0)|^2 + |x(1)|^2 + \dots + |x(N-1)|^2$  is the energy of the signal block, and  $E$  is the energy of the prediction error block.

For the autocorrelation method, a more natural approach is to use the Levinson-Durbin algorithm, [Digital Signal Processing, Principles, Algorithms and Applications, Third Edition, John G. Proakis, Dimitris G. Manolakis, Prentice Hall, International Editions, ISBN 0-13-394338-9, Chapter 11] where

$q$  is then defined according to

$$q = \left( \prod_{i=1}^p (1 - |K_i|^2) \right)^{-1} - 1, \quad (14)$$

where  $K_i$  are the reflection coefficients of the corresponding lattice filter structure obtained from the prediction polynomial, and  $p$  is the predictor order.

The ratio between highband and lowband values of  $q$  is then used to adjust the degree of spectral whitening such that the total to noise ratio of the reconstructed highband approaches that of the original highband. Here it is advantageous to control the degree of whitening utilizing the blending factor  $b$  (Eq. 6).

Assuming the total to noise ratio  $q = q_H$  is measured in the highband and  $q = q_L \geq q_H$  is measured in the lowband, a suitable choice of whitening factor  $b$  is given by the formula

$$b = 1 - \sqrt{\frac{q_H}{q_L}}. \quad (15)$$

To see this, a first step is to rewrite Eq. 6 in the form

$$A_b(z) = A(z) + (1-b)(1-A(z)). \quad (16)$$

This shows that if the signal used to estimate  $A(z)$  is filtered with the filter  $A_b(z)$ , the predicted signal is suppressed by the gain factor  $1-b$  and the prediction error is unaltered. As the total to noise ratio is the ratio of mean squared predicted signal to mean squared prediction error, a value of  $q$  prior to filtering is changed to  $(1-b)^2 q$  by the filtering operation. Applying this to the lowband signal

produces a signal with total to noise ratio  $(1-b)^2 q_L$  and under the assumption that the applied HFR method does not alter tonality, the target value  $q_H$  in the highband is reached exactly if  $b$  is chosen according to Eq. 15.

The values of  $q$  based on prediction order  $p = 2$  in each subband of a 64 channel filter bank are depicted in Fig. 5a, for the signal of Fig. 5a. Significantly higher values are reached for the harmonic part of the signal than for the noisy part. The variability of the estimates in the harmonic part is due to the chosen frequency resolution and prediction order.

#### Adaptive LPC-based whitening in the time domain

The adaptive filtering in the decoder can be done prior to, or after the high-frequency reconstruction. If the filtering is performed prior to the HFR, it needs to consider the characteristics of the HFR-method used. When a frequency selective adaptive filtering is performed, the system must deduce from what lowband region a certain highband region will originate, in order to apply the correct amount of spectral whitening to that lowband region, prior to the HFR-unit. In the example below, of a time

domain implementation of the current invention, a non-frequency selective adaptive spectral whitening is outlined. It should be obvious to any person skilled in the art that time-domain implementations of the present invention is not limited to the implementation described below.

- 5 When performing the adaptive filtering in the time domain, linear prediction using the autocorrelation method is preferred. The autocorrelation method requires windowing of the input segment used to estimate the coefficients  $a_k$ , which is not the case for the covariance method. The filter used for the spectral whitening according to the present invention is

$$H_{wp}(z, p, \rho) = 1 - \sum_{k=1}^p a_k(z\rho)^k, \quad (19)$$

- 10 where the gain factor  $O$  (in Eq. 5) is set to  $\cos$ . When the adaptive spectral whitening is performed prior to the HFR unit, an effective implementation is achieved since the adaptive filter can operate on a lower sampling rate. The lowband signal is windowed and filtered on a suitable time base with the predictor order and bandwidth expansion factors given by the encoder, according to Fig. 6. In the current implementation of the present invention the signal is low pass filtered 601 and decimated 602. 603 illustrates the adaptive filter. A window 604 is used to select the proper time segment for estimation of the  $A(z)$  polynomial, 50% overlap is used. The LPC-routine 607 extracts  $A(z)$  given the currently preferred LPC-order and bandwidth expansion factor, with a suitable relaxation. A FIR filter 608 is used to adaptively filter the signal segment. The spectrally whitened signal segments are upsampled 604, 605 and windowed together forming the input signal to the HFR unit.

#### 20 Adaptive LPC-based whitening in a subband filter bank

- The adaptive filtering can be performed effectively and robustly by using a filter bank. The linear prediction and the filtering are done independently for each of the subband signals produced by the filter bank. It is advantageous to use a filterbank where the alias components of the subband signals are suppressed. This can be achieved by e.g. oversampling the filterbank. Artifacts due to aliasing emerging from independent modifications of the subband signals, which for example adaptive filtering results in, can then be heavily reduced. The spectral whitening of the subband signals is obtained through linear prediction analogous to the time domain method described above. If the subband signals are complex valued, complex filter coefficients are used for the linear prediction as well as for the filtering. The order of the linear prediction can be kept very low since the expected number of local components in each frequency band is very small for a system with a reasonable amount of filterbank channels. In order to correspond to the same time base as the time domain LPC, the number of subband samples in each block is smaller by a factor equal to the downsampling of the filter bank. Given the low filter order and small block sizes the prediction filter coefficients are preferably obtained using the covariance method. Filter coefficient calculation and spectral whitening can be performed on a block by block basis using subband sample time step  $L$ , which is smaller than the block length  $N$ . The spectrally whitened blocks should be added together using appropriate synthesis windowing.

Feeding a maximally decimated filterbank with an input signal consisting of white gaussian noise will produce subband signals with white spectral density. Feeding an oversampled filterbank with white noise gives subband signals with coloured spectral density. This is due to the effects of the frequency responses of the analysis filters. The LPC predictors in the filterbank channels will track the filter characteristics in the case of noise-like input signals. This is an unwanted feature, and benefits from compensation. A possible solution is pre-filtering of the input signals to the linear predictors. The pre-filtering should be an inverse, or an approximation of the inverse, of the analysis filters, in order to compensate for the frequency responses of the analysis filters. The whitening filters are fed with the original subband signals, as described above. Fig. 7 illustrates the whitening process of a subband signal. The subband signal corresponding to channel  $i$  is fed to the pre-filtering block 701, and subsequently to a delay chain where the depth of the same depends on the filter order 702. The delayed signals and their conjugates 703 are fed to the linear prediction block 704, where the coefficients are calculated. The coefficients from every  $L$ -th calculation are kept by the decimator 705. The subband signals are finally filtered through the filterblock 706, where the predicted coefficients are used and updated for every  $L$ -th sample.

#### Practical implementations

The present invention can be implemented in both hardware chips and DSPs, for various kinds of systems, for storage or transmission of signals, analogue or digital, using arbitrary codecs. Fig. 8 and Fig. 9 shows a possible implementation of the present invention. In Fig. 8 the encoder side is displayed. The analogue input signal is fed to the A/D converter 801, and to an arbitrary audio coder, 802, as well as the inverse filtering level estimation unit 803, and an envelope extraction unit 804. The coded information is multiplexed into a serial bitstream, 805, and transmitted or stored. In Fig. 9 a typical decoder implementation is displayed. The serial bitstream is de-multiplexed, 901, and the envelope data is decoded, 902, i.e. the spectral envelope of the highband. The de-multiplexed source coded signal is decoded using an arbitrary audio decoder, 903. The decoded signal is fed to an arbitrary HFR unit, 904, where a highband is regenerated. The highband signal is fed to the spectral whitening unit 905, which performs the adaptive spectral whitening. Subsequently, the signal is fed to the envelope adjuster 906. The output from the envelope adjuster is combined with the decoded signal fed through a delay, 907. Finally, the digital output is converted back to an analogue waveform 908.

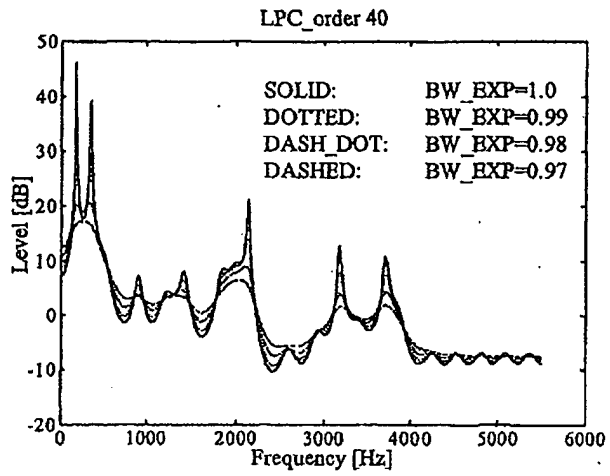
## CLAIMS

1. A method for enhancement of audio source coding systems using high-frequency reconstruction, where said source coding system comprises an encoder representing all operations performed prior to storage or transmission, and a decoder representing all operations performed after storage or transmission, characterized by:
  - at said encoder, estimating the tonal character of an original signal at a given time, and
  - at said encoder, estimating the required amount of spectral whitening at a given time, in order to obtain a similar tonal character after HFR in said decoder, given the HFR-method used in said decoder;
  - transmitting information on said amount of spectral whitening from said encoder to said decoder;
  - at said decoder, adaptively, spectrally whiten a signal prior to High Frequency Reconstruction (HFR) or after HFR, according to the spectral whitening information obtained from said encoder.
2. A method according to claim 1, characterized in that said estimation of the tonal character of the original signal is done for different frequency regions.
3. A method according to claim 1, characterized in that said estimation of the required amount of spectral whitening is done for different frequency regions.
4. A method according to claim 1, characterized in that said spectral whitening is performed in the time domain.
5. A method according to claim 1, characterized in that said spectral whitening is performed in a subband filterbank.
6. A method according to claim 1, characterized in that said estimation of required amount of spectral whitening is done by comparison of the tonal to noise signal ratios  $q$  of different subband signals obtained from subband filtering of said original signal, where said ratios are obtained using linear prediction of said subband signals.
7. A method according to claim 1, characterized in that said estimation of required amount of spectral whitening is done by comparison of the tonal to noise signal ratios  $q$  of different subband signals obtained from subband filtering of said original signal and a LFR signal, where said ratios are obtained using linear prediction of said subband signals, and said HFR signal is produced in a the same manner as said HFR in said decoder.
8. A method according to claim 1, characterized in that the amount of spectral whitening is controlled by the LPC predictor order.

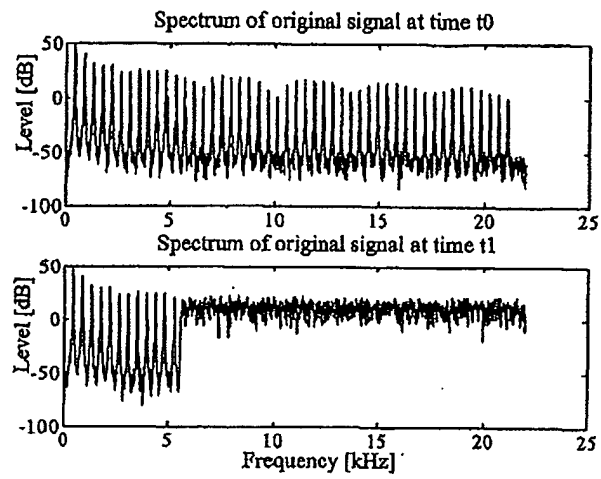


9. A method according to claim 1, characterised in that the amount of spectral whitening is controlled by the bandwidth expansion factor of the LPC polynomial.
10. A method according to claim 1, characterised in that the amount of spectral whitening is controlled by the blending factor b.
11. A method according to claim 5, characterised in that pre-filtering is included in the LPC estimation in order to compensate for the characteristic of the filterbank analysis filters.
12. An apparatus for enhancement of audio source coding systems using high-frequency reconstruction, where said source coding system comprises an encoder representing all operations performed prior to storage or transmission, and a decoder representing all operations performed after storage or transmission, characterised by:
- at said encoder, means for estimating the tonal character of an original signal at a given time,
  - at said encoder, means for estimating the required amount of spectral whitening at a given time, in order to obtain a similar tonal character after HFR in said decoder, given the HFR-method used in said decoder;
  - at said decoder, means for, adaptively, spectrally whiten a signal prior to High Frequency Reconstruction (HFR) or after HFR, according to the spectral whitening information obtained from said encoder.

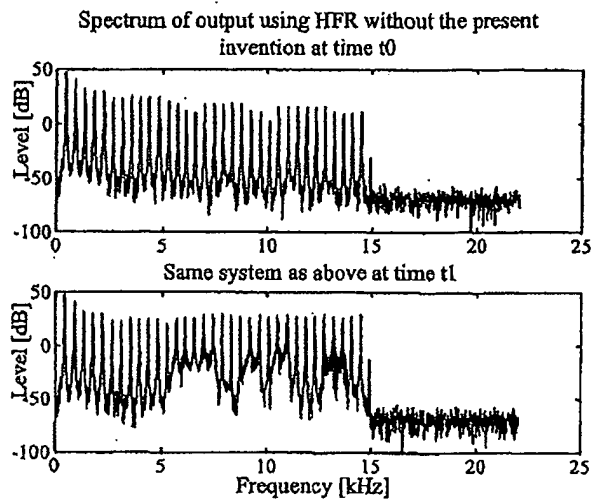
1/10

*Fig. 1*

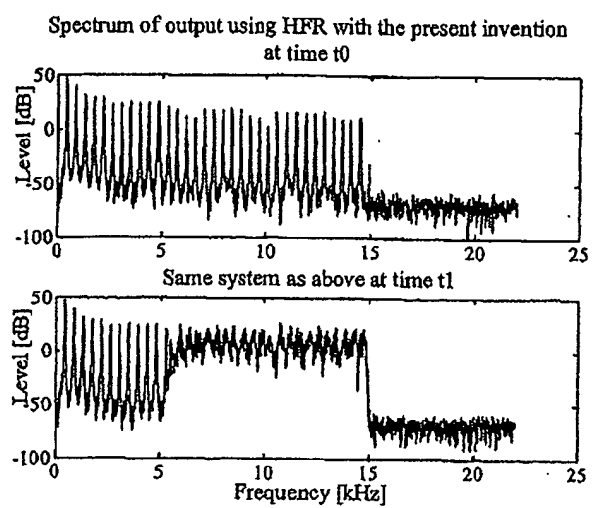
2/10

*Fig. 2*

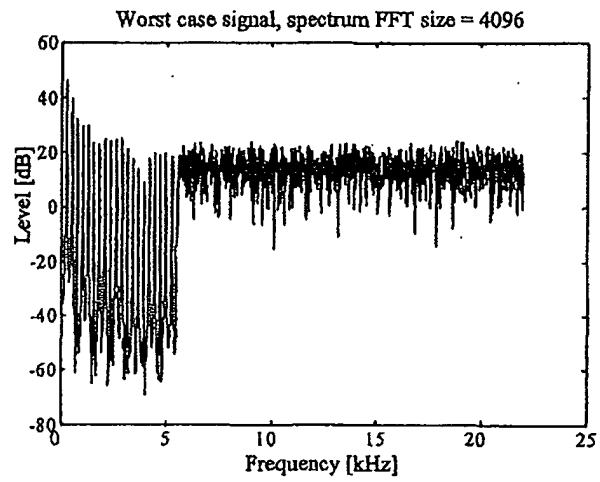
3/10

*Fig. 3*

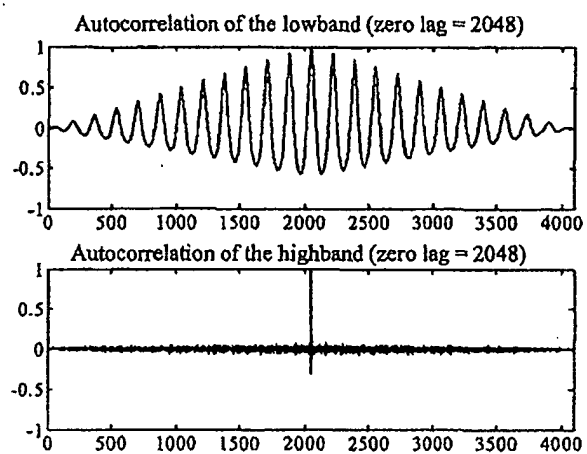
4/10

*Fig. 4*

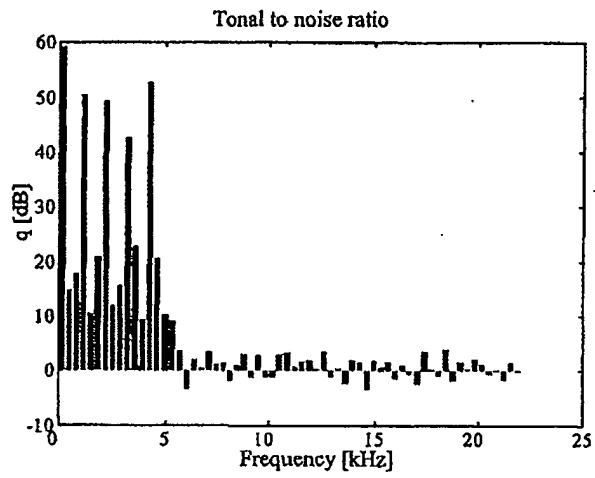
5/10

*Fig. 5a*

6/10

*Fig. 5b*

7/10

*Fig. 5c*



8/10

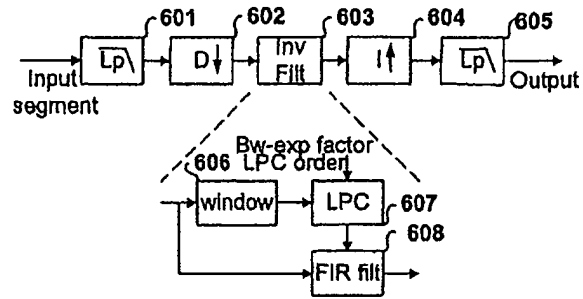


Fig. 6

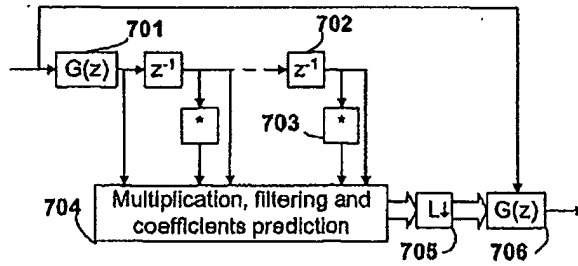
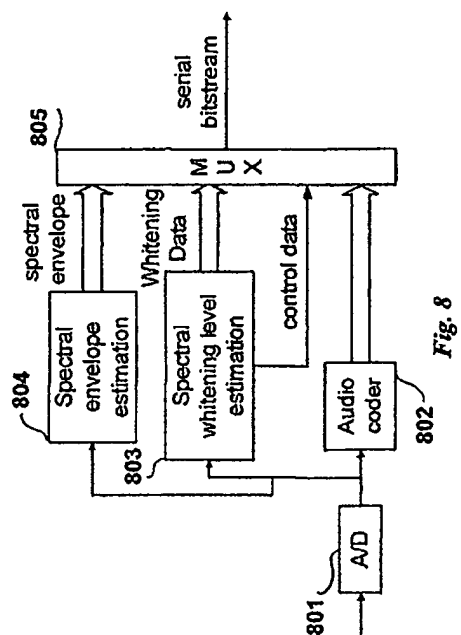


Fig. 7

9/10



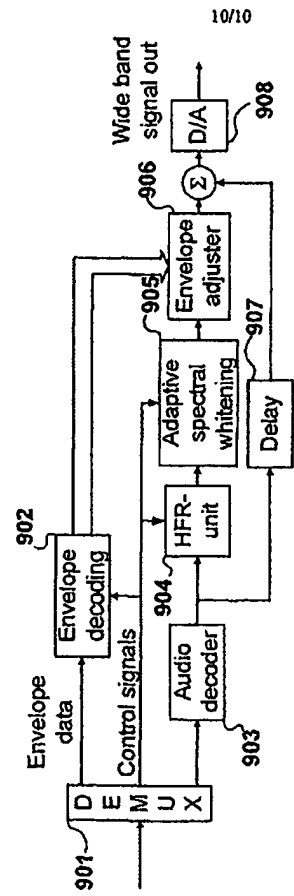


Fig. 9

【手続補正書】

【提出日】平成14年12月10日(2002.12.10)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

低帯域信号に基づく高帯域を有する高周波再生信号を生成する際に実行される高周波再生ステップの前後に信号に適応されるスペクトル白色化レベルを推定する装置であって、スペクトル白色化がスペクトル白色化フィルタを使用する周波数によって得られ、スペクトル白色化フィルタがフィルタパラメータによって適応可能である適応フィルタである装置であって、

任意の時間に符号化されるオリジナル音声信号の音色特性を推定する手段(803)であって、オリジナル音声信号が音声符号器によって符号化されて、オリジナル音声信号の低帯域のみを表す符号化音声信号が得られ、前記推定された音色特性が、符号化音声信号には含まれていない、オリジナル音声信号の高帯域の推定された音色特性を含む、手段(803)と、

推定された音色特性に基づいてスペクトル白色化フィルタの可変フィルタパラメータを決定する手段(803)と、

前記可変フィルタパラメータを前記符号化された音声信号と関連付けて、前記可変フィルタパラメータを有する前記符号化された音声信号を有するビットストリームを得る手段(805)であって、前記可変フィルタパラメータが前記符号化された音声信号に依存する、手段(805)とを備えた装置。

【請求項2】

高周波再生ステップが、実質的に低帯域の音色構造を改変せず、

前記推定手段が高帯域の音色特性以外に低帯域の音色特性も決定するように構成され、

前記決定手段が高帯域の音色と低帯域の音色特性を比較してフィルタパラメータを決定するように構成されている、請求項1記載の装置。

【請求項3】

オリジナル音声信号の低帯域で高周波再生ステップを実行して高周波再生信号を得る手段と、

前記高周波再生信号の音色特性を推定する手段とをさらに備え、

前記決定手段が前記高周波再生信号と、フィルタパラメータを決定するオリジナル音声信号の高帯域を比較するように構成されている、請求項1記載の装置。

【請求項4】

オリジナル信号の音色特性の推定が、様々な周波数領域に対して実行される、請求項1記載の装置。

【請求項5】

スペクトル白色化の必要量の推定が、様々な周波数領域に対して実行される、請求項1記載の装置。

【請求項6】

スペクトル白色化が時間ドメインで実行される、請求項1記載の装置。

【請求項7】

スペクトル白色化がサブバンドフィルタバンクで実行される、請求項1記載の装置。

【請求項8】

スペクトル白色化の必要量の推定が、オリジナル信号のサブバンド周波数から得られる様々なサブバンド信号の音色対雑音信号比の比較により行われ、前記音色対雑音信号比がサブバンド信号の線形予測を使用して得られる、請求項1記載の装置。

## 【請求項 9】

スペクトル白色化の必要量の推定が、オリジナル信号及び高周波復元信号のサブバンドろ波から得られる様々なサブバンド信号の音色対雑音信号比の比較により行われ、前記音色対雑音信号比がサブバンド信号の線形予測を使用して得られ、前記高周波復元信号がデコードにおける高周波復元信号と同様に生成される、請求項 1 記載の装置。

## 【請求項 10】

前記スペクトル白色化フィルタが L P C 多項式を得るための線形予測により得られるフィルタ係数を有するフィルタであり、前記フィルタパラメータが、L P C 多項式の予測子次数、L P C 多項式の帯域幅膨張係数、又はろ波された信号及び未処理対照物の混合量を示す混合係数を示す、請求項 1 記載の装置。

## 【請求項 11】

オリジナル音声信号の低帯域を表す符号化音声信号の復号化バージョンに基づいて出力信号を生成する装置であって、前記符号化音声信号がスペクトル白色化フィルタの可変フィルタパラメータと関連付けられ、前記可変フィルタパラメータが任意の時間でのオリジナル音声信号の高帯域の音色特性に依存している装置であって、前記符号化音声信号と関連付けられる前記可変フィルタパラメータを得る手段 (901) と、

前記符号化音声信号の復号化バージョンで高周波再生ステップを実行して、高周波再生信号を生成する高周波再生部 (904) と、

前記復号化バージョン又は前記高周波再生信号をろ波する適応スペクトル白色化フィルタ (905) とを備え、

前記適応スペクトル白色化フィルタが可変パラメータを有し、前記可変パラメータが前記符号化音声信号と関連付けられる前記可変フィルタパラメータに従って設定される、装置。

## 【請求項 12】

フィルタバンク解析フィルタの特性を補償するために、事前ろ波が線形予測コード化推定に含まれる、請求項 11 記載の装置。

## 【請求項 13】

前記適応スペクトル白色化フィルタが、

ろ波される信号をウィンドウ処理する手段 (606) と、

ウィンドウ処理された信号の L P C 多項式を得る L P C 手段 (607) であって、任意の時間、可変フィルタパラメータとして L P C 次数及び帯域幅膨張係数に応答する L P C 手段 (607) と、

ろ波される信号をろ波する F I R フィルタであって、前記 L P C 手段によって得られる L P C 多項式により設定される F I R フィルタとを備える、請求項 11 の装置。

## 【請求項 14】

低帯域信号に基づく高帯域を有する高周波再生信号を生成する際に実行される高周波再生ステップの前後に信号に適応されるスペクトル白色化レベルを推定する方法であって、スペクトル白色化がスペクトル白色化フィルタを使用するろ波によって得られ、前記スペクトル白色化フィルタがフィルタパラメータによって適応可能な適応フィルタである方法であって、

任意の時間に符号化されるオリジナル音声信号の音色特性を推定するステップであって、オリジナル音声信号が音声符号器によって符号化されて、オリジナル音声信号の低帯域のみを表す符号化音声信号が得られ、前記推定された音色特性が、前記符号化音声信号には含まれない、オリジナル音声信号の高帯域の推定された音色特性を含む、ステップと、前記推定された音色特性に基づいて前記スペクトル白色化フィルタの可変フィルタパラメータを決定するステップと、

前記可変フィルタパラメータを前記符号化された音声信号と関連付けて、前記可変フィルタパラメータを有する前記符号化された音声信号を有するビットストリームを得るステップであって、前記可変フィルタパラメータが前記符号化された音声信号に依存しているス

テップとを備える方法。

【請求項 15】

オリジナル音声信号の低帯域を表す符号化音声信号の復号化バージョンに基づいて出力信号を生成する方法であって、前記符号化音声信号がスペクトル白色化フィルタの可変フィルタパラメータと関連付けられ、前記可変フィルタパラメータが任意の時間でオリジナル音声信号の高帯域の音色特性に依存する方法であって、  
前記符号化音声信号と関連付けられた前記可変フィルタパラメータを得るステップと、  
前記符号化音声信号の復号化バージョンで高周波再生ステップを実行して、高周波再生信号を生成するステップと、  
白色化フィルタ（905）を使用して前記復号化バージョン又は前記高周波再生信号をろ波するステップとを備え、  
前記適応スペクトル白色化フィルタが可変パラメータを有し、前記可変パラメータが符号化音声信号と関連付けられた前記可変フィルタパラメータに従って設定される、方法。

【請求項 16】

オリジナル音声信号を符号化してその符号化バージョンを得るエンコーダであって、  
請求項 1 に係るスペクトル白色化レベルを推定する装置と、  
オリジナル音声信号を符号化して、その符号化バージョンを得る音声エンコーダ（802）と、  
オリジナル音声信号のスペクトルエンベロープを推定して、推定されたスペクトルエンベロープを得る手段（804）と、  
ビットストリームを取得するために、オリジナル音声信号の前記符号化バージョン、前記スペクトル白色化フィルタのフィルタパラメータ、及び推定されたスペクトルエンベロープを多重化するマルチプレкса（805）とを備えるエンコーダ。

【請求項 17】

オリジナル音声信号の符号化バージョン、推定されたスペクトルエンベロープ、及びスペクトル白色化フィルタに適用されるフィルタパラメータを含むビットストリームを復号するデコーダであって、  
前記オリジナル音声信号の符号化バージョン、前記推定されたスペクトルエンベロープ、及び前記フィルタパラメータを抽出するビットストリームデマルチプレкса（901）と、  
前記オリジナル音声信号の符号化バージョンを復号して低帯域信号を得る音声デコーダ（903）と、  
前記推定されたスペクトルエンベロープを復号するエンベロープデコーダと、  
請求項 11 に係る、出力信号を生成する装置と、  
適応可能にスペクトル白色化された高周波再生信号及び前記復号された音声信号の遅延バージョンを加算して広帯域出力信号を取得する加算器とを備えたデコーダ。

【請求項 18】

オリジナル音声信号を符号化してその符号化バージョンを得る方法であって、  
請求項 14 に係るスペクトル白色化レベルを推定するステップと、  
オリジナル音声信号を符号化（802）し、その符号化バージョンを得るステップと、  
オリジナル音声信号のスペクトルエンベロープを推定（804）し、推定されたスペクトルエンベロープを得るステップと、  
ビットストリームを取得するために、前記オリジナル音声信号の符号化バージョン、前記スペクトル白色化フィルタのフィルタパラメータ、及び前記推定されたスペクトルエンベロープを多重化（805）するステップと備えた方法。

【請求項 19】

オリジナル音声信号の符号化バージョン、推定されたスペクトルエンベロープ、及びスペクトル白色化フィルタに適用されるフィルタパラメータを含むビットストリームを復号する方法であって、  
前記オリジナル音声信号の符号化バージョン、前記推定されたスペクトルエンベロープ、

及び前記フィルタパラメータを抽出（９０１）し、  
前記オリジナル音声信号の符号化バージョンを復号（９０３）して低帯域信号を取得し、  
前記推定されたスペクトルエンベロープを復号し、  
請求項１５に従って出力信号を生成し、  
適応可能にスペクトル白色化された高周波再生信号及び前記復号された音声信号の遅延バージョンを加算して広帯域出力信号を取得する方法。

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PC7/SE 01/02510
<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
IPC7: G10L 19/00, H03H 7/30 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Magnetic documentation searched (classification system followed by classification symbols)		
IPC7: G10L, H03H		
Documents searched other than magnetic documentation to the extent that such documents are included in the fields searched		
SE, DK, FI, NO classes as above		
Electronic data base searched during the international search (names of data base and, where practicable, search terms used)		
MPI-DATA, INSPEC		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Criteria of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J. Acoust. Soc. Am., Volume 66, No 6, December 1979, John Makhoul et al., "Predictive and residual encoding of speech", page 1633 - page 1641, abstract; section V	1-12
A	Signal Processing VII Theories and Applications, Proceedings of EUSIPCO-94, Seventh European Signal Processing Conference, Edinburgh, Scotland, U.K., 13-16 September 1994, Vol. II, Holger Carl et al; "Bandwidth Enhancement of Narrow-Band Speech Signals", pages 1178-1181, see abstract; section 3	1-12
A	NO 8603872 A1 (SITE LABORATORIES INCORPORATED), 3 July 1986 (03.07.86), claims 1-6	1-12
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family search.		
<p>* Special categories of cited documents</p> <p>"A" documents defining the general state of the art which is not considered to be of particular relevance</p> <p>"B" documents defining the general state of the art which is not considered to be of particular relevance</p> <p>"C" documents which may have priority claims or which are cited to establish the priority date of another claim or to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"D" documents which are cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"E" documents published prior to the international filing date but later than the priority date of claim</p> <p>"F" documents published after the international filing date or priority date and not in accordance with the provisions laid down in the preamble of the present treaty</p> <p>"G" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"H" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"I" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"J" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"K" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"L" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"M" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"N" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"O" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"P" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"Q" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"R" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"S" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"T" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"U" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"V" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"W" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"X" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"Y" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p> <p>"Z" documents of particular relevance the document is cited to establish the priority date of another claim or to establish the priority date of another claim</p>		
Date of the actual completion of the international search		Date of mailing of the international search report
19 February 2002		27-02-2002
Name and mailing address of the ISA/ Swedish Patent Office Box 5059, S-102 02 STOCKHOLM Telephone No. +46 8 666 02 56		Authorized officer Peder Gjerwald saeter/AE Telephone No. +46 8 782 25 00

Form PC7/PA/210 (revised sheet) (July 1993)



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/SE 01/02510

C (Continued) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category <sup>1</sup>	Citation of document, with indications, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 0345379 A2 (LILJERYD, LARS, GUSTAF), 3 August 2000 (03.08.00), abstract --	1-12
A	WO 9857436 A2 (LILJERYD, LARS, GUSTAF), 17 December 1998 (17.12.98), abstract --	1-12
A	US 5915235 A (ANDREW P. DEJACO ET AL.), 22 June 1999 (22.06.99), column 3, line 66 - column 4, line 13, claim 6B -----	1, 12

Form PCT/ISA/210 (continuation of record sheet) (July 1995)

INTERNATIONAL SEARCH REPORT  
Information on patent family members

28/01/02

International application No.  
PCT/SE 01/02510

Patent documents cited in search report	Publication date	Patent family member(s)	Publication date
WO 8603872 A1	03/07/86	CA 1239701 A	26/07/88
		DE 3587251 A,T	13/05/93
		EP 0208712 A,B	21/01/87
		US 4914701 A	03/04/90
		WO 8603873 A	03/07/86
		CA 1301337 A	19/05/92
		US 4790016 A	06/12/88
WO 0045379 A2	03/08/00	AU 2585600 A	18/08/00
		AU 2585700 A	18/08/00
		AU 3283999 A	15/09/99
		BR 0009138 A	27/11/01
		EP 1064739 A	03/01/01
		EP 1157374 A	28/11/01
		SE 9900256 D	00/00/00
		SE 9903552 D	00/00/00
		SE 9903553 D	00/00/00
		WO 0045378 A	03/08/00
WO 9857436 A2	17/12/98	AU 5684298 A	07/08/98
		AU 7446598 A	30/12/98
		BR 9805989 A	31/08/99
		CN 1272259 T	01/11/00
		EP 0940015 A	08/09/99
		JP 2001521648 T	06/11/01
		SE 512719 C	02/05/00
		SE 9702213 D	00/00/00
		SE 9800256 A	11/12/98
		SE 9704634 D	00/00/00
		DE 19880227 C,T	01/02/01
US 5916235 A	22/06/99	NONE	

Form PCT/ISA/210 (patent family member) (July 1998)

---

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TC),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GI,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(72)発明者 フレドリック・ヘン

スウェーデン国 S-16831 ブロンマ リタルヴェーゲン 14

(72)発明者 ラルス・ヴィレモエス

スウェーデン国 S-17556 ユエルフェーラ マンドリンヴェーゲン 22

【公報種別】特許法第17条の2の規定による補正の掲載  
 【部門区分】第6部門第2区分  
 【発行日】平成18年1月5日(2006.1.5)

【公表番号】特表2004-514179(P2004-514179A)  
 【公表日】平成16年5月13日(2004.5.13)  
 【年通号数】公開・登録公報2004-018  
 【出願番号】特願2002-543427(P2002-543427)  
 【国際特許分類】

G 1 0 L 19/02 (2006.01)  
 G 1 0 L 13/00 (2006.01)  
 G 1 0 L 19/00 (2006.01)

【F I】

G 1 0 L 7/04 F  
 G 1 0 L 7/04 E  
 G 1 0 L 7/02 D  
 G 1 0 L 9/08 J

【誤訳訂正書】  
 【提出日】平成17年9月20日(2005.9.20)  
 【誤訳訂正1】  
 【訂正対象書類名】明細書  
 【訂正対象項目名】0028  
 【訂正方法】変更  
 【訂正の内容】  
 【0028】

64チャネルフィルタバンクの各サブバンドでの予測次数 $p=2$ に基づく $q$ の値は、図5aの信号について図5cに示す。雑音部分より信号の高調波部分について大幅に高い値に達する。高調波部分での推定値の可変性は、選ばれる周波数分解能及び予測次数による。

【誤訳訂正2】  
 【訂正対象書類名】明細書  
 【訂正対象項目名】0029  
 【訂正方法】変更  
 【訂正の内容】  
 【0029】

(時間ドメインにおける適応LPCベースの白色化)

デコーダでの適応ろ波は、高周波復元の前後に実行できる。ろ波がHFRの前に実行される場合、使用されるHFR法の特徴を考慮する必要がある。周波数選択適応ろ波が実行される場合は、低帯域領域に正しい量のスペクトル白色化を適用するために、システムは、HFR部の前で、特定の低帯域領域が生じる低帯域領域から控除しなければならない。本発明の時間ドメイン法の以下の例では、非周波数選択適応スペクトル白色化が概略的に説明される。本発明の時間ドメイン法が後述の実施例に制限されないことはこの技術に精通した者にとって明らかである。

【誤訳訂正3】  
 【訂正対象書類名】明細書  
 【訂正対象項目名】0031  
 【訂正方法】変更  
 【訂正の内容】  
 【0031】

(サブバンドフィルタバンクでの適応LPCベースの白色化)

適応ろ波は、フィルタバンクを使用して効果的かつ堅牢に実行できる。線形予測及びろ波は、フィルタバンクにより発生するサブバンド信号のそれぞれに、独立して実行される。サブバンド信号のエイリアス成分が抑制されるフィルタバンクを使用することが有利である。これは、例えばフィルタバンクをオーバーサンプルすることによって達成できる。例えば適応ろ波の結果生じる、サブバンド信号の個々の変型から出現するエイリアスのためのアーチファクトは、大きく削減できる。サブバンド信号のスペクトル白色化は、前述された時間ドメイン法に類似する線形予測を通して取得される。サブバンド信号が複素値である場合、ろ波だけではなく線形予測にも複素数のフィルタ係数が使用される。各周波数バンドの音色成分の予想数は妥当な量のフィルタバンクチャネルのあるシステムに対しては非常に小さいため、線形予測の次数を非常に低く保つことができる。時間ドメインLPCと同じ時間ベースに対応するために、各ブロックのサブバンドサンプル数はフィルタバンクのダウンサンプルに等しい係数分小さくなる。低フィルタ次数及び小さなブロックサイズを考慮すると、予測フィルタ係数は、好ましくは共分散法を使用して得られる。フィルタ係数の算出及びスペクトル白色化は、ブロック長 $N$ より小さいサブバンドサンプルタイムステップ $L$ を使用してブロック単位で実行できる。スペクトル白色化されたブロックは、適切な合成ウィンドウ処理を使用してともに追加されなければならない。

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第2区分

【発行日】平成18年11月9日(2006.11.9)

【公表番号】特表2004-514179(P2004-514179A)

【公表日】平成16年5月13日(2004.5.13)

【年通号数】公開・登録公報2004-018

【出願番号】特願2002-543427(P2002-543427)

【国際特許分類】

G 1 0 L 19/02 (2006.01)

G 1 0 L 13/00 (2006.01)

G 1 0 L 19/00 (2006.01)

【F I】

G 1 0 L 7/04 F

G 1 0 L 7/04 E

G 1 0 L 7/02 D

G 1 0 L 9/08 J

【誤訳訂正書】

【提出日】平成18年9月13日(2006.9.13)

【誤訳訂正1】

【訂正対象書類名】明細書

【訂正対象項目名】0002

【訂正方法】変更

【訂正の内容】

【0002】

(発明の背景)

高帯域が低帯域から外挿される音声信号の高周波復元において、HFRシステムで一般に使用されるような粗いエンベロープ調整で達成可能であるよりもさらに広い範囲まで復元された高帯域の音声成分を制御する手段を有することが重要である。これが必要な理由は、声等の多くの音声信号及び多くの音響機器の音色成分が、通常は、高周波領域と比較して低周波領域(すなわち4~5kHz以下)においてより強いためである。極端な例は、低帯域が非常にはっきりした調波系列であり、高帯域が多かれ少なかれ純水な雑音の場合である。これに取り組む1つの方法は、再構築された高帯域に雑音を適応可能に追加することである(適応雑音追加[PCT/SE00/00159号])。しかしながら、これは低帯域の音色の特性を抑制するほど十分ではない場合があり、再構築された高帯域に反復性の「バズ音」特性を与える。さらに、雑音の正しい時間特性を達成することが困難である場合がある。一方が高い調波密度(低ピッチ)をもち、他方が低い調波密度(高ピッチ)をもつ2つの調波系列が混合されるときに別の問題が生じる。高ピッチの調波系列が低帯域では他方を支配するが、高帯域では支配しない場合、HFRは高ピッチ信号の調波に高帯域を支配させ、復元された高帯域音をオリジナルに比べて「金属的」にする。前述のシナリオのいずれも、HFRシステムで一般的に使用されるエンベロープ調整を使用して制御することはできない。HFR信号のスペクトルエンベロープ調整の間に一定程度のスペクトル白色化が導入される実例もある。これは、その特定の程度のスペクトル白色化が所望の場合には満足の行く結果を与えるが、その特定の程度のスペクトル白色化から恩恵を受けない信号の抜粋には深刻なアーチファクトを生じさせる。

【誤訳訂正2】

【訂正対象書類名】明細書

【訂正対象項目名】0003

【訂正方法】変更

【訂正の内容】

【0003】

(発明の概要)

本発明は、HFR方法に共通して生じる「バズ音」及び「金属音」の問題に関する。本発明は、エンコーダ側で高性能な検出アルゴリズムを使用して、デコーダで適用されるスペクトル白色化の好適な量を推定する。スペクトル白色化は、周波数のみならず経時的にも変化し、複製される高帯域の調波含有量を制御するための最善の手段を確実にする。本発明は、サブバンドフィルタバンク法においてだけでなく、時間ドメイン法においても実行可能である。

【誤訳訂正3】

【訂正対象書類名】明細書

【訂正対象項目名】0015

【訂正方法】変更

【訂正の内容】

【0015】

1つの方法は、スペクトル白色化の適切な量を推定するために自己相関を使用する。検出器は、ソース範囲（すなわち、HFRがデコーダ内で基にする周波数範囲）及びターゲット範囲（すなわち、デコーダ内で復元される周波数範囲）の自己相関関数を推定する。図5aに、低帯域の調波系列及び高帯域の白色雑音をもつ、最悪の場合の信号を記述する。図5bに、様々な自己相関関数を示す。ここで、低帯域はきわめて相互に関連付けられるが、高帯域は相互に関連付けられていないことが明らかである。最小の遅延より大きな遅延について最大相関が高帯域及び低帯域に得られる。この2つの比率が、デコーダで適用されるスペクトル白色化の最適な度合を算出するために使用される。本発明を前述において概略したように実現すると、相関の算出にFFTを使用することが好適である場合がある。数列  $x(n)$  の自己相関は次式により定義される。

【数8】

$$r_{xx}(m) = FFT^{-1}(|X(k)|^2) \quad (8)$$

式中、

【数9】

$$X(k) = FFT(x(n)) \quad (9)$$

である。

【誤訳訂正4】

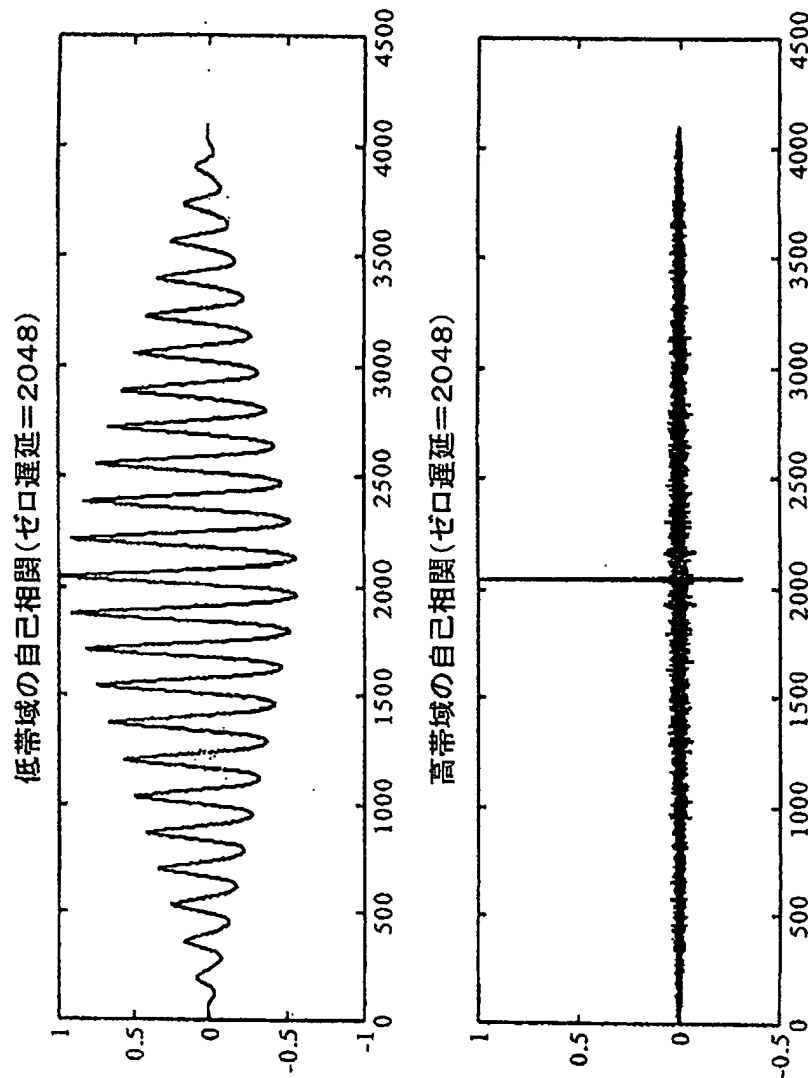
【訂正対象書類名】図面

【訂正対象項目名】図5b

【訂正方法】変更

【訂正の内容】

【図 5 b】



【誤訳訂正 5】

【訂正対象書類名】明細書

【訂正対象項目名】0028

【訂正方法】変更

【訂正の内容】

【0028】

64チャネルフィルタバンクの各サブバンドでの予測次数  $p=2$  に基づく  $q$  の値は、図 5 a の信号について図 5 c に示す。信号の調波部分では雑音部分よりも大幅に高い値が得られる。調波部分での推定値の可変性は、選ばれる周波数分解能及び予測次数による。

【誤訳訂正 6】

【訂正対象書類名】明細書

【訂正対象項目名】0029

【訂正方法】変更



## 【訂正の内容】

【0029】

(時間ドメインにおける適応LPCベースの白色化)

デコーダでの適応ろ波は、高周波復元の前後に実行できる。ろ波がHFRの前に実行される場合、使用されるHFR法の特徴を考慮する必要がある。周波数選択適応ろ波が実行される場合は、低帯域領域に正しい量のスペクトル白色化を適用するために、システムは、HFR部の前で、特定の高帯域領域がどのような低帯域領域から生じるものであるかを決定しなければならない。本発明の時間ドメイン法の以下の例では、非周波数選択適応スペクトル白色化が概略的に説明される。本発明の時間ドメイン法が後述の実施例に制限されないことはこの技術に精通した者にとって明らかである。

## 【誤訳訂正7】

【訂正対象書類名】明細書

【訂正対象項目名】0030

【訂正方法】変更

【訂正の内容】

【0030】

時間ドメインで適応ろ波を実行する際は、自己相関法を使用する線形予測が好ましい。自己相関法は、係数 $\alpha_k$ を推定するために使用される入力セグメントのウィンドウ処理を必要とするが、このことは共分散法には当てはまらない。本発明によるスペクトル白色化に使用されるフィルタは、

【数17】

$$H_{mv}(z, p, \rho) = 1 - \sum_{k=1}^P \alpha_k (z\rho)^{-k} \quad (17)$$

であり、式中、利得係数G(数式5)が1に設定される。適応スペクトル白色化がHFR部の前に実行される場合、適応フィルタはさらに低いサンプリング速度で動作できるため、効果的な実行が達成される。低帯域信号は、図6に従って、エンコーダにより与えられた予測子次数及び帯域幅膨張係数をもってウィンドウ処理され、適切な時間ベースでろ波される。本発明のこの実施例では、信号は低域ろ波され(601)、大量に除去される(602)。603は適応フィルタを示す。ウィンドウ606はA(z)多項式の推定のための適切な時間セグメントを選択するために使用され、50%の重複が使用される。LPCルーチン607は、最新に選ばれたLPC次数及び帯域幅膨張係数が与えられたA(z)を、適切な緩和をもって抽出する。FIRフィルタ608は、信号セグメントを適応可能にろ波するために使用される。スペクトル白色化された信号セグメントはアップサンプルされ(604、605)、共にウィンドウ処理され、HFR部に対する入力信号を形成する。

## 【誤訳訂正8】

【訂正対象書類名】明細書

【訂正対象項目名】請求項1

【訂正方法】変更

【訂正の内容】

【請求項1】 低帯域信号に基づく高帯域を有する高周波再生信号を生成する際に実行される高周波再生ステップの前後に信号に適応されるスペクトル白色化レベルを推定する装置であって、スペクトル白色化がスペクトル白色化フィルタを使用するろ波によって得られ、スペクトル白色化フィルタがフィルタパラメータによって適応可能である適応フィルタである装置であって、

任意の時間に符号化されるオリジナル音声信号の音色特性を推定する手段(803)であって、オリジナル音声信号が音声符号器によって符号化されて、オリジナル音声信号の低帯域のみを表す符号化音声信号が得られ、前記推定された音色特性が、符号化音声信号には含まれていない、オリジナル音声信号の高帯域の推定された音色特性を含む、手段(

803)と、

推定された音色特性に基づいてスペクトル白色化フィルタの可変フィルタパラメータを決定する手段(803)と、

前記可変フィルタパラメータを前記符号化された音声信号と結合して、前記可変フィルタパラメータを有する前記符号化された音声信号を有するビットストリームを得る手段(805)であって、前記可変フィルタパラメータが前記符号化された音声信号に依存する手段(805)とを備えた装置。

【誤訳訂正9】

【訂正対象書類名】明細書

【訂正対象項目名】請求項11

【訂正方法】変更

【訂正の内容】

【請求項11】 オリジナル音声信号の低帯域を表す符号化音声信号の復号化バージョンに基づいて出力信号を生成する装置であって、前記符号化音声信号がスペクトル白色化フィルタの可変フィルタパラメータと結合され、前記可変フィルタパラメータが任意の時間でのオリジナル音声信号の高帯域の音色特性に依存している装置であって、

前記符号化音声信号と結合される前記可変フィルタパラメータを得る手段(901)と

、前記符号化音声信号の復号化バージョンで高周波再生ステップを実行して、高周波再生信号を生成する高周波再生部(904)と、

前記復号化バージョン又は前記高周波再生信号をろ波する適応スペクトル白色化フィルタ(905)とを備え、

前記適応スペクトル白色化フィルタが可変パラメータを有し、前記可変パラメータが前記符号化音声信号と結合される前記可変フィルタパラメータに従って設定される、装置。

【誤訳訂正10】

【訂正対象書類名】明細書

【訂正対象項目名】請求項14

【訂正方法】変更

【訂正の内容】

【請求項14】 低帯域信号に基づく高帯域を有する高周波再生信号を生成する際に実行される高周波再生ステップの前後に信号に適応されるスペクトル白色化レベルを推定する方法であって、スペクトル白色化がスペクトル白色化フィルタを使用するろ波によって得られ、前記スペクトル白色化フィルタがフィルタパラメータによって適応可能である適応フィルタである方法であって、

任意の時間に符号化されるオリジナル音声信号の音色特性を推定するステップであって、オリジナル音声信号が音声符号器によって符号化されて、オリジナル音声信号の低帯域のみを表す符号化音声信号が得られ、前記推定された音色特性が、前記符号化音声信号には含まれていない、オリジナル音声信号の高帯域の推定された音色特性を含む、ステップと、

前記推定された音色特性に基づいて前記スペクトル白色化フィルタの可変フィルタパラメータを決定するステップと、

前記可変フィルタパラメータを前記符号化された音声信号と結合して、前記可変フィルタパラメータを有する前記符号化された音声信号を有するビットストリームを得るステップであって、前記可変フィルタパラメータが前記符号化された音声信号に依存しているステップとを備える方法。

【誤訳訂正11】

【訂正対象書類名】明細書

【訂正対象項目名】請求項15

【訂正方法】変更

【訂正の内容】

【請求項15】 オリジナル音声信号の低帯域を表す符号化音声信号の復号化バージョンに基づいて出力信号を生成する方法であって、前記符号化音声信号がスペクトル白色化フィルタの可変フィルタパラメータと結合され、前記可変フィルタパラメータが任意の時間でオリジナル音声信号の高帯域の音色特性に依存する方法であって、

前記符号化音声信号と結合される前記可変フィルタパラメータを得るステップと、

前記符号化音声信号と結合された前記可変フィルタパラメータを得るステップと、

前記符号化音声信号の復号化バージョンで高周波再生ステップを実行して、高周波再生信号を生成するステップと、

白色化フィルタ（905）を使用して前記復号化バージョン又は前記高周波再生信号をろ波するステップとを備え、

前記適応スペクトル白色化フィルタが可変パラメータを有し、前記可変パラメータが符号化音声信号と結合された前記可変フィルタパラメータに従って設定される、方法。

【公報種別】特許法第17条の2の規定による補正の掲載  
 【部門区分】第6部門第2区分  
 【発行日】平成19年4月12日(2007.4.12)

【公表番号】特表2004-514179(P2004-514179A)  
 【公表日】平成16年5月13日(2004.5.13)  
 【年通号数】公開・登録公報2004-018  
 【出願番号】特願2002-543427(P2002-543427)  
 【国際特許分類】

G 1 0 L 19/02 (2006.01)  
 G 1 0 L 13/00 (2006.01)  
 G 1 0 L 19/00 (2006.01)

【F I】

G 1 0 L 7/04 F  
 G 1 0 L 7/04 E  
 G 1 0 L 7/02 D  
 G 1 0 L 9/08 J

【誤訳訂正書】  
 【提出日】平成19年2月20日(2007.2.20)  
 【誤訳訂正1】  
 【訂正対象書類名】明細書  
 【訂正対象項目名】0014  
 【訂正方法】変更  
 【訂正の内容】  
 【0014】

(エンコーダ側の検出器)

本発明では、現在使用されているHFR方法を考慮して、可能な限りオリジナルに類似した高帯域を得るために、デコーダで使用されるスペクトル白色化(LPC次数、帯域幅膨張係数又は混合係数)の最良の度合を査定するためにエンコーダ側の検出器を使用する。デコーダで使用されるスペクトル白色化の度合の適切な推定値を得るにはいくつかの方法が使用可能である。後述する説明では、HFRアルゴリズムは、高周波の生成の間に低帯域スペクトルの音色構造を実質的には改変しない、つまり生成された高帯域は低帯域と同じ音色の特性を備えると仮定する。このような仮定を立てることができない場合、以下の検出は合成による解析を使用して、つまり、オリジナル信号の低帯域及び高帯域に関する比較研究を実行するよりも、むしろ、エンコーダでオリジナル信号にHFRを実行し、高周波再生信号と、フィルタパラメータを決定するオリジナル音声信号との高帯域を比較することにより行われる。

【誤訳訂正2】  
 【訂正対象書類名】明細書  
 【訂正対象項目名】請求項4  
 【訂正方法】変更  
 【訂正の内容】

【請求項4】 オリジナル信号の音色特性の推定が、異なった周波数領域に対して実行される、請求項1記載の装置。

【誤訳訂正3】  
 【訂正対象書類名】明細書  
 【訂正対象項目名】請求項5  
 【訂正方法】変更  
 【訂正の内容】

【請求項5】 スペクトル白色化の必要量の推定が、異なった周波数領域に対して実行される、請求項1記載の装置。

【誤訳訂正4】

【訂正対象書類名】 明細書

【訂正対象項目名】 請求項8

【訂正方法】 変更

【訂正の内容】

【請求項8】 スペクトル白色化の必要量の推定が、オリジナル信号のサブバンドろ波から得られる異なったサブバンド信号の音色対雑音信号比の比較により行われ、前記音色対雑音信号比がサブバンド信号の線形予測を使用して得られる、請求項1記載の装置。

【誤訳訂正5】

【訂正対象書類名】 明細書

【訂正対象項目名】 請求項9

【訂正方法】 変更

【訂正の内容】

【請求項9】 スペクトル白色化の必要量の推定が、オリジナル信号のサブバンドろ波から得られる異なったサブバンド信号の音色対雑音信号比の比較により行われ、前記音色対雑音信号比がサブバンド信号の線形予測を使用して得られる、請求項1記載の装置。

【誤訳訂正6】

【訂正対象書類名】 明細書

【訂正対象項目名】 0002

【訂正方法】 変更

【訂正の内容】

【0002】

(発明の背景)

高帯域が低帯域から外挿される音声信号の高周波復元において、HFRシステムで一般に使用されるような粗いエンベロープ調整で達成可能であるよりもさらに広い範囲まで復元された高帯域の音声成分を制御する手段を有することが重要である。これが必要な理由は、声等の多くの音声信号及び多くの音響機器の音色成分が、通常は、高周波領域と比較して低周波領域（すなわち4～5 kHz以下）においてより強いためである。極端な例は、低帯域が非常にはっきりした調波系列であり、高帯域が多かれ少なかれ純粋な雑音の場合である。これに取り組む1つの方法は、再構築された高帯域に雑音を適応可能に追加することである（適応雑音追加[PCT/SE00/00159号]）。しかしながら、これは低帯域の音色の特性を抑制するほど十分ではない場合があり、再構築された高帯域に反復性の「バズ音」特性を与える。さらに、雑音の正しい時間特性を達成することが困難である場合がある。一方が高い調波密度（低ピッチ）をもち、他方が低い調波密度（高ピッチ）をもつ2つの調波系列が混合されるときに別の問題が生じる。高ピッチの調波系列が低帯域では他方を支配するが、高帯域では支配しない場合、HFRは高ピッチ信号の調波に高帯域を支配させ、復元された高帯域音をオリジナルに比べて「金属的」にする。前述のシナリオのいずれも、HFRシステムで一般的に使用されるエンベロープ調整を使用して制御することはできない。HFR信号のスペクトルエンベロープ調整の間に一定程度のスペクトル白色化が導入される実例もある。これは、その特定の程度のスペクトル白色化が所望の場合には満足の行く結果を与えるが、その特定の程度のスペクトル白色化から恩恵を受けない信号の抜粋には深刻なアーチファクトを生じさせる。